

Neuvermessung der Datenökonomie

herausgegeben von Thomas Straubhaar

EDITION HWWI

Hamburg University Press

Neuvermessung der Datenökonomie
herausgegeben von Thomas Straubhaar



Reihe Edition HWWI
herausgegeben von Thomas Straubhaar
Band 6

Neuvermessung der Datenökonomie



herausgegeben von Thomas Straubhaar

Hamburg University Press
Verlag der Staats- und Universitätsbibliothek Hamburg
Carl von Ossietzky

Impressum

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen National- bibliografie; detaillierte bibliografische Daten sind im Internet über https://portal.dnb.de abrufbar.



ISSN 1865-7974

ONLINE-AUSGABE
Die Online-Ausgabe dieses Werkes ist eine Open-Access-Publikation und ist auf den Verlagswebseiten frei verfügbar. Die Deutsche Nationalbibliothek hat die Online-Ausgabe archiviert. Diese ist dauerhaft auf dem Archivserver der Deutschen Nationalbibliothek (https://portal.dnb.de) verfügbar.

DOI https://doi.org/10.15460/HUP.HWWI.6.212

ISBN
Print: 978-3-943423-91-4
EPUB: 978-3-943423-94-5

Satz Hamburg University Press

Covergestaltung Hamburg University Press unter Verwendung eines Fotos von Free-Photos auf Pixabay (https://pixabay.com/images/id-768432)

Druck und Bindung Books on Demand (Norderstedt)

Verlag
Hamburg University Press, Verlag der Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky, Hamburg (Deutschland), 2021
https://hup.sub.uni-hamburg.de

Inhalt

Vorwort

Was taugen Verfahren der Vergangenheit, um die Ökonomie der Zukunft zu vermessen? Der Urmeter der Ökonomie, das Bruttoinlandsprodukt (BIP), dürfte für kommende Zeiten kaum mehr das Maß aller Dinge sein. Zu stark sind die durch technologische Fortschritte verursachten grundsätzlichen Mängel des BIP-Konzepts. Nicht erst seitdem das Coronavirus viele alte Gesetzmäßigkeiten in Frage stellt, bestehen Zweifel, inwieweit Messverfahren aus der Hochzeit der Industrialisierung dem strukturellen Wandel zunächst zu einer Dienstleistungsgesellschaft, später zu einer Digitalwirtschaft und nun zu einer Datenökonomie gerecht werden (können). Vielmehr drängte sich bereits vorher an mancher Stelle eine Neuorientierung auf.

Die nachfolgenden Überlegungen und Erkenntnisse wollen zu einer erfolgreichen Suche nach neuen Verfahren der Messung ökonomischer Aktivitäten beitragen. Dabei geht es im ersten Teil um Messprobleme und darum, inwieweit sie bei der Diagnose und Prognose wirtschaftlicher Entwicklungen zu analytischen Fehlern führen. Der zweite Teil vermittelt einen Überblick über neue empirische Ansätze, mit deren Hilfe besser erfasst werden kann, was sich in der (Daten-)Ökonomie abspielt.

Präsentiert werden Einsichten und Ergebnisse des Forschungsprojekts „Neuvermessung der Weltwirtschaft“, das von der NORDAKADEMIE-Stiftung (Elmshorn/Hamburg) finanziell gefördert wurde. Für diese großzügige Unterstützung sei allen Verantwortlichen herzlich gedankt.

Ebenso ist der Universität Hamburg zu danken, die dem Forschungsprojekt Gastrecht gewährte und die für einen Großteil der Autorinnen und Autoren eine stimulierende und kreative Arbeitsatmosphäre schuf.

Hamburg, im Sommer 2021
Thomas Straubhaar
Projektleiter und Herausgeber

Einleitung

Datenwirtschaft: Was ist neu und anders?

Thomas Straubhaar

„The world's most valuable resource is no longer oil, but data." So eindeutig fällt das Urteil der britischen Wochenzeitschrift *The Economist* aus.[1] Wie richtig die These ist, lässt sich mit einer Rangliste der weltweit wertvollsten Firmen veranschaulichen. Mitte 2021 steht Apple mit einem Börsenwert (BW) von rund 2,0 Billionen Euro an der Spitze.[2] Insgesamt sieben der zehn höchstbewerteten Unternehmen haben mit (digitaler Infrastruktur und Plattformen) der Datenökonomie zu tun. Neben Apple sind das Microsoft (BW 1,8 Bio. Euro), Amazon (BW 1,6 Bio. Euro), die Google-Mutter Alphabet (BW 1,3 Bio. Euro) und Facebook (BW 0,7 Bio. Euro) sowie die chinesischen (Daten-)Handelsplattformen Tencent und Alibaba.[3]

Daten sind der Rohstoff der Digitalisierung. Sie sind, was Öl und Gas für das Industriezeitalter waren. Sie treiben in nahezu allen ökonomischen Bereichen weitreichende Verhaltensänderungen an – auch im Alltag. Sie verändern Informationswege und Kommunikationskanäle. Sie ermöglichen neue Formen von Wertschöpfung und Arbeit, Produktion und Konsum. Kein Lebensbereich, der nicht durch die Datenwirtschaft radikal betroffen wäre. Das gilt auch für ökonomische Theorie und Empirie.

---

[1] Economist (2017).

[2] Finanzen100 (2021).

[3] Finanzen100 (2021). Allerdings steht mit Saudi Aramco ein Unternehmen der traditionellen Rohstoffökonomie auf Platz zwei der Rangliste. Die Erdölfördergesellschaft aus Saudi-Arabien war im Dezember 2019 mit dem größten Börsengang in der Weltgeschichte sofort mit Platz eins ins Ranking eingestiegen (siehe dazu: Tagesschau 2021). Im Laufe von 2020 fiel jedoch der Börsenwert von Saudi Aramco auf 1,5 Bio. Euro.

Die Coronapandemie war ein Augenöffner, was Digitalisierung und Datenwirtschaft in der Praxis bedeuten. Der Bevölkerung wurde nahezu über Nacht flächendeckend vorgeführt, wie sehr digitale Technologien bisherige Gewohnheiten umzukrempeln vermögen. Onlinebestellungen machten das Einkaufen in Supermärkten überflüssig. In Schulen und Universitäten trat Fernunterricht an die Stelle von Präsenzlehre. Homeoffice führte zu leeren Geschäftsbüros. Videobesprechungen erübrigten Fernreisen. An vielen Stellen ersetzte ein Datentransfer den Warenhandel. Der ökonomische Strukturwandel von einer Güter- zu einer Dienstleistungs- und weiter zu einer Digitalisierungswirtschaft wurde im Laufe der Coronapandemie und ihrer Bekämpfung noch einmal beschleunigt.

Die gravierenden Unterschiede zwischen Güter- und Datenökonomie erfordern neue Konzepte, um das Wesen von Big Data zu erfassen, zu vergleichen und zu verstehen. Eine neue Datenökonomik für eine neue Datenökonomie ist das eine.[4] Eine Neuvermessung der Wirtschaft ist das andere.[5] Beides zusammen wird notwendig, wenn neben einer sichtbaren Waren- eine unsichtbare Datenwirtschaft wichtig(er) wird. Oder wie es Liran Einav und Jonathan Levin von der Stanford University ausdrücken: „More granular and comprehensive data surely allow improved measurements of economic effects and outcomes, better answers to old questions, and help in posing new questions and enabling novel research designs. We also believe that new data may change the way economists approach empirical research, as well as the statistical tools they employ."[6]

Grundlagen einer Datenökonomie

Die Datenökonomie unterscheidet sich fundamental von der Güterwirtschaft.[7] Denn anders als Güter sind Daten nicht dinglicher Natur. Sie sind unsichtbar. Sie wabern virtuell losgelöst von Standorten, In- und Ausland, durch Internet und Orbit. Big Data ist „huge in volume, … high in velocity, … diverse in variety,

[4] Stellvertretend für andere Einav/Levin (2014).

[5] Stellvertretend für andere Harding/Hersh (2018).

[6] Einav/Levin (2014).

[7] Stellvertretend für andere Wambach/Müller (2018), IMF (2019) und Azkan et al. (2020) aus volkswirtschaftlicher sowie Ochs et al. (2019) aus sozioökonomischer Sicht.

... exhaustive in scope, ... fine-grained in resolution, ... relational in nature, ... flexible in scalabilty".[8]

Loslösung ist das Wesen der Datenwirtschaft. Ökonomische Aktivitäten entbinden und entgrenzen sich gleichzeitig von Dingen, Fabriken, Standorten oder Wirtschaftsräumen. Sie werden „entdinglicht", wenn Daten anstelle von Gütern produziert, gehandelt und verkauft werden – also beispielsweise Musik nicht mehr auf Schallplatten, sondern über Spotify angeboten wird. Sie werden „entterritorialisiert", wenn mehr und mehr Glieder der Wertschöpfung in virtuelle Wolken („Clouds") des Orbits verlagert werden und beispielsweise Professorinnen und Professoren nicht mehr in Hörsälen unterrichten, sondern Studierende die Angebote virtueller Online-Universitäten nutzen. Und sie werden „entstaatlicht", weil das Internet keine physischen Grenzen und damit auch keine Landesgrenzen, Volkswirtschaften oder Nationalökonomien mehr kennt.

Daten machen die Welt zum Dorf. Raum und Territorialität verlieren an Bedeutung. In der Datenwirtschaft ist eine geografische Festlegung von Herkunft, Verarbeitung, Vermarktung und Nutzung schwierig bis kaum möglich. Information und Kommunikation machen an nationalen Grenzen nicht halt. Alles und jedes ist in Echtzeit überall bekannt und verfügbar. Staatsgrenzen verschwimmen und Unternehmensgrenzen verschwinden.

Daten zu transferieren und auszutauschen kostet nahezu nichts. Sie sind um Dimensionen rascher hin und her zu bewegen als Güter, aber auch als Dienstleistungen, bei denen oft die Präsenz der anbietenden und/oder nachfragenden Person erforderlich ist – beispielsweise bei medizinischen Operationen. Digitalisierung und Datenökonomie ermöglichen entpersonalisierte Dienstleistungen. Beispielsweise erlaubt die Telemedizin, klinische Eingriffe auf Distanz durchzuführen – Patient und Ärztin müssen sich nicht vor Ort treffen. Binäre Daten (insbesondere in Text-, Wort- oder Bildformaten) ersetzen Warenhandel oder persönliche Kontakte.

Die Datenökonomie ist eine Ökonomie der Transparenz. „Gläsern" zu werden schafft nicht nur Nachteile. Es sorgt auch dafür, dass Informationen überall und jederzeit für alle gleichermaßen zur freien Nutzung bereitstehen. Noch nie in der Menschheitsgeschichte war es auch nur annähernd so einfach und umfassend möglich, so viel Information und Wissen so schnell und so billig weltweit und in Echtzeit auszutauschen wie heute. Das hat die Lebensqualität nahezu überall enorm verbessert. Angebote und Preise, Löhne und Zinsen lassen

8 Saetnan et al. (2018), S. 3.

sich transparent vergleichen, was Wettbewerb fördert und Verbraucherrechte verbessert. Genauso konnten Herrschaftswissen, Informationsmonopole und Wissensdefizite beseitigt und damit demokratische Bewegungen und die Rechte von Minderheiten gestärkt werden.

Datenökonomie wird eine neue Dezentralität ermöglichen, die mit der Industrialisierung in weiten Teilen verloren gegangen war. In der Güterwirtschaft erzwangen die immensen Fixkosten der industriellen Produktion die Zentralisierung der Produktion in riesigen Fabriken, in denen Massen von Beschäftigten aktiv tätig waren. Von (angestrebt nur) einem Standort aus wurden dann möglichst hohe Stückzahlen hergestellt. So ließen sich Durchschnittskosten senken.

Mit der Zentralisierung ging eine Spezialisierung einher. Die Glieder der industriellen Wertschöpfungskette wurden in immer kleinere Teile zerlegt. So lohnte es sich für jedes einzelne Glied, spezielle Facharbeiter auszubilden und ebenso spezielle maschinelle Fabrikationshilfen anzuschaffen. Je größer die Stückzahlen waren, umso stärker konnten die Spezialisierungsvorteile ausgenutzt und umso breiter die Spezialisierungsinvestitionen (für Anschaffung von Wissen und Maschinen) umgelegt werden. In Folge dessen sanken die durchschnittlichen Kosten pro Stück weiter.

In der Post-Corona-Datenwirtschaft wird „hybrid“ zum Zauberwort werden. Es bezeichnet die intelligente Mischung zwischen Vereinheitlichung (um die Vorteile der Standardisierung, also der „Economies of Scale“ zu nutzen) und Vielfalt (um individuell unterschiedlichen Ansprüchen gerecht zu werden). Mit hybriden Strukturen wird eine kluge Aufteilung von Zentralität und Dezentralität bei Produktion und Konsum einhergehen. Vieles wird an einer Stelle einheitlich programmiert, vorgegeben, organisiert und umgesetzt. Manches jedoch wird vor Ort, zu Hause oder beim Kunden geleistet, hergestellt oder ausgedruckt.

Über allem und jedem, was in einer „hybriden“ Wirtschaft und Gesellschaft passiert und geschieht, wird das Damoklesschwert von Optimierung und Verbesserung schweben. Es wird nahezu automatisch und ansonsten mit Anreizen – vielleicht auch Sanktionen – dafür sorgen, dass in der Datenwirtschaft des Post-Corona-Zeitalters mehr als jemals zuvor Effizienz und Effektivität zum Maß aller Dinge werden. Das hilft, ökonomischen Leerlauf zu verringern und ökologische Verschwendung zu vermeiden.

In der Datenwirtschaft wird künstliche Intelligenz neben menschliches Wissen treten und den Alltag mehr und mehr erleichtern und verbessern. Sobald Menschen digital miteinander kommunizieren und arbeiten, online einkaufen und tauschen, betreten sie das Spielfeld, auf dem Algorithmen und

selbstlernende Systeme das Sagen haben und die Regeln bestimmen. Dort wird alles und jedes stets und immerzu gemessen, bewertet und verglichen. Was aktuell das Beste ist, wird zum Maßstab für alles andere. Fehler und Abweichungen werden sofort erkannt und sogleich behoben.

Optimierung, Standardisierung und Skalierung bilden die entscheidungs- und handlungsleitende Trilogie der Datenökonomie. Was sich irgendwo bewährt (hat), wird in gleicher oder ähnlicher Weise überall kopiert. Zunächst einzelne, im Laufe der Zeit jedoch immer mehr Glieder der Wertschöpfung und später auch Bereiche des gesellschaftlichen Miteinanders werden sich dem steten Qualitätstest zu stellen haben. Was passt, überlebt, alles andere geht unter.

Grundlagen einer Datenökonomik

Die Datenwirtschaft verlangt ein radikales Umdenken in nahezu allen Dimensionen ökonomischen Denkens, das sich stark an physischen Gütern orientiert, mit abnehmendem Grenznutzen beim Konsum und steigenden Grenzkosten bei der Herstellung.

Bei Daten gibt es – wiederum anders als bei Gütern – keine Abnutzung. Zwischen verschiedenen Nutzungszwecken von Daten besteht kein Konflikt. „Nur, weil jemand bestimmte Daten nutzt, schließt dies – anders als eben beim Öl – nicht die parallele oder sequenzielle Nutzung derselben Daten durch andere aus.“[9] Während viele Personen die gleichen Daten gleichzeitig auf verschiedene Art und Weise nutzen, wird die Qualität der aus den Daten gewonnenen Informationen und Einsichten nicht schlechter, sondern vielleicht sogar besser. Gerade die Weiterverwendung, die Verknüpfung und Mehrfachnutzung von Daten erzeugt positive Netzwerkeffekte und macht deren Gebrauch für alle Beteiligten eher noch wertvoller. Wer Daten teilt, hat (oft) nicht weniger, sondern mehr davon. Aus der gemeinsamen Datennutzung kann ein Mehrwert für alle entstehen – etwa, wenn aus einer Vielzahl unterschiedlicher Blogeinträge ein Allgemeinwissen darüber entsteht, wie Alltagsprobleme – beispielsweise die Entfernung von Kleiderflecken – einfach behoben werden können.

Daten sind also eher ein öffentliches als ein privates Gut. Wenn die Fixkosten für Netzinfrastruktur, Hard- und Software einmal entstanden sind, können Daten mit sehr geringen Grenzkosten, also nahezu ohne Zusatzkosten, von

---

9 Haucap (2018), S. 472.

vielen anderen verarbeitet, verbreitet und weiterverwendet werden. Ein Ausschluss von Nutzern durch Bezahlschranken hat wenig mit technischen Kapazitätsgrenzen, Abnutzungseffekten oder gegenseitiger Beeinträchtigung zu tun. Er erfolgt aus kommerziellen Gründen, um die Hersteller für ihren Aufwand zu entschädigen. Aber nur für einen Bruchteil der Datentransfers kommen (bis anhin) Bezahlsysteme zur Anwendung. Die meisten Daten jedoch sind mehr oder weniger frei verfügbar, so wie Allgemeinwissen auch allen ohne direkte Gegenleistung zugänglich ist.

Die Datenökonomie trägt den Keim einer Monopolisierung in sich. Big-Data-Konzerne können eine Marktmacht erlangen, die es ihnen ermöglicht, „Big Profits" zulasten der Verbraucherinnen und Verbraucher zu erzielen. Ursache hierfür sind die immensen Fixkosten der Datenverarbeitung auf der einen und die nahezu vollständige Abwesenheit variabler Kosten der Datennutzung auf der anderen Seite. Derartige Skaleneffekte sind der Humus, auf dem Monopole blühen.[10]

Hinzu kommen Pioniervorteile und Pfadabhängigkeiten, wie sie für digitale Plattformen charakteristisch sind. Erstere entstehen, weil Regeln und Standards derjenige setzen kann, wer zuerst kommt. Letztere spielen eine Rolle, weil sich Nutzer an gewisse Marken und Produkte, bestimmte Verfahren, Abläufe oder Verhaltensweisen so sehr gewöhnt haben, dass Kundentreue und Gewohnheiten dazu führen, immer wieder bekannte Produkte zu kaufen und zu nutzen. Plattformen haben so viele Daten und Informationen über ihre Nutzer, dass sie mit Hilfe von Algorithmen präzise Kundenprofile erstellen und damit den gläsernen Nutzer schaffen können.

Pioniere haben einen Startvorsprung. Sie sind oft in einer Monopolstellung. So sind sie in der Lage, zu Beginn neuer Produktzyklen jene Gewinne zu erwirtschaften, die später notwendig sind, um sich die im Laufe der Zeit als Nachahmer auf den Markt drängende Konkurrenz vom Halse zu halten. Dank einer dominanten Marktposition, bestehender Kundenbeziehungen und früher erzielter Gewinnpolster können sich einmal etablierte Big-Data-Firmen jederzeit auf einen Preiswettbewerb einlassen, der für sie zwar schmerzlich, für Konkurrenten jedoch ruinös ist.

Laufende Profite wiederum sind die eigene Kapitalquelle (oder die für außenstehende Eigenkapitalgeber entscheidende Attraktion, neues Wagniskapital einzuschießen), aus der die Big-Tech-Firmen die gewaltigen Investitionen in

---

[10] Einschränkend dazu Dewenter/Lüth (2016).

stetige Innovationen finanzieren können, die mit dem Aufbau einer sich verstetigenden Monopolposition einhergehen. Damit aber ist vorgezeichnet, dass „Big Data“ mit „Big Profits“ einhergehen kann, was erklärt, wieso die großen Datenkonzerne die Weltrangliste des Big Business (gemessen am Wert der Börsenkapitalisierung) anführen.

Warum eine Neuvermessung der Wirtschaft erforderlich wird

Was sich in der Ökonomie an dynamischem Wandel abspielt, verlangt nach einer Neuvermessung der wirtschaftlichen Aktivitäten, insbesondere der Wertschöpfung und der Produktivität. Denn die Ökonomik findet ihre (statistischen) Grundlagen in der Agrarwirtschaft und deren Weiterentwicklung im Zeitalter der Industrialisierung. [11] Sie hat sich bereits mit Dienstleistungswirtschaft und Wissensgesellschaft schwergetan. Beide entzogen sich gängigen Verfahren der Messung, des Zählens und Bewertens, wie sie in der Güterwirtschaft mit ihren haptischen, physisch fassbaren Produkten charakteristisch sind. Gleiches und noch viel mehr anderes gilt in besonderem Maße für die Datenwirtschaft mit ihren nicht greifbaren, in virtuellen Räumen fließenden Bits und Bytes.[12]

Herkömmliche (makro-)ökonomische Messverfahren wie das BIP sind völlig ungeeignet, das Wesen der Datenökonomie und den Charakter der Daten abzubilden. Das BIP ist ein „statistisches Konstrukt“ und kein „hartes Faktum“.[13] Es ist eher ein errechnetes als ein gemessenes Maß. Es misst Quantitäten, die sich

[11] Die lange Entstehungsgeschichte des (makro-)ökonomischen Vermessungswesens begann im 17. Jahrhundert mit William Pettys ersten Schätzungen für des Volkseinkommens in England. Der Hofarzt Ludwigs XV., François Quesnay, im 18. Jahrhundert, Karl Marx im 19. Jahrhundert sowie Léon Walras und Alfred Marshall im Übergang zum 20. Jahrhundert führten sie weiter. Die „Volkswirtschaftlichen Gesamtrechnungen“ erhielten mit der Entwicklung eines BIP-Konzepts in der Zwischenkriegszeit einen Schub. Die damals gewaltigen makroökonomischen Verwerfungen machten eine bessere Vermessung ökonomischer Aktivitäten erforderlich. Aus kontroversen Ideen von John Maynard Keynes, Richard Stone, James Meade, Simon Kuznets und Wassily Leontief entstand in den 1950er-Jahren ein international standardisiertes Verfahren, mit dessen Hilfe das BIP weltweit nach einheitlichen Methoden gemessen wurde. Vgl. dazu Stobbe (1988) sowie als Übersicht OECD (2014).

[12] Vgl. dazu exempalrisch Varian (2014) oder Garboden (2020).

[13] Die Unterscheidung zwischen „harten Fakten“, die sich „objektiv“ durch „zählen“, „messen“ oder „wiegen“ erfassen lassen, und „statistischen Konstrukten“, die eher „subjektiv“ durch „bewerten“ und „beobachten“ errechnet werden, geht auf den Wirtschaftsnobelpreisträger Richard Stone (2013) zurück, der zwischen „primary facts“ und „empirical constructs“ unterscheidet. „Primary Facts“ entsprechen der direkt messbaren „Quantität“, während „statistical constructs“ eher die „Qualität“ widerspiegeln, die sich indirekt aus der subjektiven Bewertung ergibt.

mehr oder weniger exakt und „objektiv“ erfassen lassen. Deren Qualität ergibt sich aus der subjektiven Wertschätzung von Herstellern und Kunden und aus Angebot und Nachfrage durch Kosten und Preise. In die BIP-Messung fließen dann die Umsätze (also faktisch das Produkt aus Mengen und Preisen) ein, von denen lediglich die Vorleistungen subtrahiert werden, um die Bruttowertschöpfung zu erhalten.[14]

Stärken und Schwächen des BIP-Konzepts sind hinlänglich bekannt.[15] Viele fundamentale konzeptionelle Mess-, Erfassungs- und Bewertungsprobleme haben immer schon zu harscher Kritik an einem BIP-Fetischismus der Ökonomik geführt.[16] Entsprechend der begrenzten Aussagekraft des BIP-Konzepts gab und gibt es seit je eine Vielzahl von Ansätzen zur Verbesserung.[17] Die größte internationale Beachtung erlangte der seit 1990 veröffentlichte Human Development Index (HDI) der Vereinten Nationen.[18] Der HDI ergänzt und verknüpft das BIP mit der Lebenserwartung bei der Geburt und der Dauer der Schulbildung und will so den sozioökonomischen Entwicklungsstand innerhalb von Ländern messen und vergleichen.

---

[14] Alle Varianten der Erfassung – beispielsweise über Kosten (wie Löhne und Zinsen) – sind nur „künstlich“ vorgenommene indirekte Berechnungen (und eben keine direkten Messungen), die aus den ursprünglichen Umsätzen durch Umformungen und Plausibilitätsrechnungen abgeleitet beziehungsweise geschätzt werden.

[15] Angesichts der Fülle von Publikationen, die sich mit den Vor- und Nachteilen, Stärken und Schwächen sowie grundsätzlichen Mängeln beschäftigen, sei hier nur auf Stobbe (1994), den „Klassiker“ der deutschsprachigen Literatur, und auf eine Zusammenfassung von Hübl (2007) hingewiesen. Räth und Braakmann (2014) präsentieren einen umfassenden, kompetenten Überblick zur Generalrevision der Volkswirtschaftlichen Gesamtrechnungen 2014 für den Zeitraum 1991 bis 2014.

[16] Die jüngere Kritik begann mit den Publikationen des Club of Rome (Meadows et al. 1972) und ist unverändert im Gange bei den heutigen Degrowth-Diskussionen (Schmelzer 2016, Konzeptwerk Neue Ökonomie 2017).

[17] Dazu gehört beispielsweise, dass nach Vorgaben der Europäischen Kommission in den BIP-Berechnungen nun auch illegale Transaktionen wie Drogenhandel, Alkohol- und Tabakschmuggel oder Bestechungsgelder erfasst werden (Taschowsky 2015). Anders als in anderen EU-Ländern war in Deutschland die „Bruttowertschöpfung“ der Prostitution schon vorher im BIP miteinbezogen – und vom Statistischen Bundesamt für das Jahr 2013 auf 7,3 Mrd. Euro geschätzt worden, was knapp einem drittel Prozent des deutschen BIP entspricht (Plickert 2014). Schneider und Boockmann (2020), S. 22 schätzen, dass die Wertschöpfung in der sogenannten Schattenwirtschaft (definiert als ökonomische Aktivitäten, aus denen Einkommen erzielt, aber staatliche Regulierung, Besteuerung oder Erfassung vermieden wird) im Jahr 2020 für Deutschland 322 Mrd. Euro erreicht, was etwa 9 % des offiziell ausgewiesenen BIP entspricht. Allerdings beinhaltet diese Schätzung zu einem kleinen Teil auch bereits Aktivitäten, die im offiziellen BIP erfasst werden (wie beispielsweise den Drogenhandel). „Rechnet man diese heraus, liegt die Wertschöpfung (der Schattenwirtschaft T. S.) bei schätzungsweise 210 Mrd. Euro“ oder rund 6 % des offiziellen BIP (Enste 2019), S. 153.

[18] UNDP (2020).

Eine Reihe anderer monetärer Wohlfahrtsindikatoren (die das BIP um zusätzliche soziale und ökologische Parameter ergänzen) wie der Index of Sustainable Economic Welfare (ISEW) und der daraus entstandene Genuine Progress Indicator (GPI) oder der Nationale Wohlfahrtsindex (NWI) genauso wie Mehrkomponentenindikatoren wie der Index of Economic Well-Being (IEW-B) und der Index of Social Health (ISH) wurden als Alternativen zum BIP entwickelt.[19]

In Deutschland suchte die Enquete-Kommission „Wachstum, Wohlstand, Lebensqualität – Wege zu nachhaltigem Wirtschaften und gesellschaftlichem Fortschritt in der Sozialen Marktwirtschaft" nach Messkonzepten jenseits des BIP.[20] Ziel der vom Deutschen Bundestag eingesetzten insgesamt 62 Mitglieder, die mehr als zwei Jahre lang als parlamentarische Expertinnen und Experten und externe Sachverständige in der Enquete-Kommission mitgewirkt haben, war es, „in Ergänzung zum BIP" einen „ganzheitlichen Wohlstands- beziehungsweise Fortschrittsindikator" zu schaffen. Der neue Urmeter der Ökonomie solle insbesondere „den materiellen Lebensstandard, Zugang zu und Qualität von Arbeit, die gesellschaftliche Verteilung von Wohlstand, die soziale Inklusion und Kohäsion, intakte Umwelt und Verfügbarkeit begrenzter natürlicher Ressourcen, Bildungschancen und Bildungsniveaus, Gesundheit und Lebenserwartung, Qualität öffentlicher Daseinsvorsorge, sozialer Sicherung und politischer Teilhabe sowie die subjektiv von den Menschen erfahrene Lebensqualität und die Zufriedenheit" beachten.[21]

Als Ergebnis ihres 844-seitigen Schlussberichts veröffentlichte die Enquete-Kommission im Mai 2013 ein „erweitertes BIP", das neben dem Materiellen auch die Wohlstandsdimensionen Soziales/Teilhabe und Ökologie abbildet. „Die aus zehn zentralen Variablen bestehenden W³ Indikatoren sollen künftig darüber Auskunft geben, wie es in Deutschland um Wohlstand und Lebensqualität steht."[22] Fakt ist, dass mit der Vorlage des Schlussberichts auch gleich das Ende eines Neuanfangs besiegelt wurde. Denn seit 2013 wurden das „erweiterte BIP" und die zehn Leitindikatoren weder faktisch in den statistischen Alltag umgesetzt noch in der wirtschaftspolitischen Arena weiterverfolgt.

---

[19] Die Enquete-Kommission (2013) liefert in ihrem Kapitel 6 (S. 302–327) eine ausführliche Stärken-Schwächen-Analyse zu zahlreichen anerkannten und aussagekräftigen Wohlstandsmaßen und in Kapitel 7 (S. 327–331) empirische Erfahrungen mit deren konkreter Anwendung in vier Ländern (Australien, Kanada, Großbritannien und USA).

[20] Enquete-Kommission (2013).

[21] Enquete-Kommission (2013), S. 231.

[22] Enquete-Kommission (2013), S. 28.

So bleiben die enormen Anstrengungen der statistischen Ämter zu würdigen, Teilaspekte der Volkswirtschaftlichen Gesamtrechnungen (VGR) weiterzuentwickeln und zum Beispiel Aspekte der Einkommens- und Vermögensverteilung, Nettogrößen, Haushaltsperspektive und andere neue Themen einzubeziehen. „Allerdings besteht hier die Gefahr einer Überfrachtung der VGR und die Ableitung statistischer Artefakte. Das gegenwärtige System der VGR weist bereits eine hohe Komplexität auf. ... Auch für die Volkswirtschaftlichen Gesamtrechnungen gibt es eine Grenze der Universalität.“[23] Diese Feststellung dürfte in Zeiten starker technologischer Umwälzungen und der daraus folgenden enormen Dynamik von Veränderungen in allen Bereichen von Wirtschaft, Gesellschaft und Politik ganz besonders zutreffen.

Wenn ökonomische Aktivitäten losgelöst von physischen Gütern, Erdboden und Nationalstaaten in der Datenwirtschaft virtueller Netzwelten erfolgen, hat das gewaltige Rückwirkungen auf die Ökonomie (die Wirtschaft als Objekt, das analysiert wird) und die Ökonomik (die Wirtschaftswissenschaft als Subjekt, das analysiert). Löst sich die Wirtschaft von Räumen, Personen, Firmen und Volkswirtschaften, lässt sich (digitale) Wertschöpfung nicht mehr verlässlich fixen Quellen oder einzelnen Volkswirtschaften zuordnen. Wenn im Orbit unsichtbare Daten anstelle dinglicher Produkte gehandelt werden, dann wirken makroökonomische Analysen altbacken und überholt, die Menschen, Unternehmen, Sektoren, Regionen oder gar Staaten in Gruppen zusammenfassen und als homogene Untersuchungsgegenstände handhaben.

Die Wertschöpfung im Internet, der virtuelle Handel mit digitalen Daten sowie die Effekte einer Sharing Economy, in der gerade „teure“ langlebige Güter – wie Wohnungen, Autos oder Elektrogeräte – gemeinsam genutzt und nicht einzeln erworben werden, entziehen sich in beachtlichen Teilen der sachlichen Erfassung, räumlichen Zuordnung und zeitlichen Abgrenzung. Deshalb sind das BIP und seine Messverfahren von geringerer Aussagekraft denn je.

Das Ende der Volkswirtschaft als homogen abgrenzbarer Untersuchungsgegenstand und die am Anfang stehende Entwicklung hin zu digitalen Handlungsebenen verlangt neue Perspektiven der Ökonomik auf das, was sich in der Ökonomie abspielt. Alte Strukturen werden in Frage gestellt, die das wirtschaftswissenschaftliche Denken, die statistische Vermessung der Wirtschaft und das wirtschaftspolitische Handeln in den letzten Jahrhunderten bestimmt haben.

---

[23] Räth (2016), S. 113.

„GDP is a statistic designed for mass production. … It falls short in measuring intangible benefits.“[24] Und noch weniger wird das BIP in der Lage sein, die digitale Ökonomie abzubilden. Entsprechend größer wird die Kluft zwischen BIP-Messung und Realität. Und umso größer ist die Gefahr, wirtschaftspolitische Fehlschlüsse zu ziehen.

Statistische Ämter haben in der Vergangenheit (durchaus erfolgreich) eine Vielzahl von Schätzverfahren, Behelfsrechnungen und Korrekturmaßnahmen angewandt, um die Aussagekraft und Belastbarkeit des BIP auch in Zeiten dynamischer Innovationen zu sichern. Aber bereits den Strukturwandel von der „dinglichen“ Industrie- zur „unsichtbaren“ Dienstleistungs- und Wissensgesellschaft konnte das BIP-Konzept nur noch begrenzt abbilden. „Für spezielle Analysen neuerer Phänomene wie Digitalisierung, Globalisierung oder ‚Sharing Economy‘ sind jedoch ergänzende Messansätze erforderlich.“[25]

Eine neue Datenökonomik soll helfen, besser zu verstehen, was sich im Zeitalter der Daten, der Algorithmen, der künstlichen Intelligenz, der selbstlernenden Automaten und der Automobilität selbstständiger Fahrzeuge alles ändert und welche Chancen und Risiken sich daraus ableiten lassen. Einen kleinen Beitrag zum besseren Verständnis einer großen Wissenslücke zu liefern, ist Ziel der in diesem Buch vereinten Beiträge, die sich in zwei Themenfelder teilen.

Inhaltsüberblick

Im ersten Teil geht es um eine alte Angst, die neues Interesse findet. Wie bereits vor hundert Jahren beschäftigt besorgte Zeitgenossen das Paradox, dass große technologische Innovationen die Welt nicht voranbringen, sondern zurückwerfen. Damals sagten führende Intellektuelle der Weltwirtschaft eine „säkulare Stagnation“ voraus, obwohl immense Fortschritte im Gesundheitswesen, bei Verkehr und Transport, Kommunikation, Logistik und Mobilität erkennbar

---

[24] Coyle (2014), S. 125.

[25] Statistisches Bundesamt (2017), S. 36. Der in Theorie und erster Anwendung wohl am weitesten entwickelte Ansatz ist das GDP-B-Konzept von Brynjolfsson et al. (2019). Das GDP-B geht über das BIP hinaus („beyond GDP“), weil es von der Nutzenseite her versucht, den Wert (Benefits) frei verfügbarer Daten der Digitalwirtschaft zu messen. Beispielsweise versuchen die Autoren zu schätzen, welchen monetären Betrag Nutzerinnen und Nutzer für Facebook zu zahlen bereit wären. Aus ähnlichen Experimenten wird dann ein gesamtwirtschaftlicher Mehrwert ermittelt, der sich aus Digitalwirtschaft und Datenökonmie ergibt und um den das BIP zu vergrößern wäre. Allein Facebook würde demgemäß das US-BIP um 0,05 bis 0,11 % pro Jahr höher ausfallen lassen, als es tatsächlich ausgewiesen wird.

wurden. Der Menschheit stehe eine lange Phase ökonomischen Nullwachstums bevor.

Ähnliche Prognosen eines verlangsamten wirtschaftlichen Fortschritts finden sich heutzutage erneut. Vielfach wird das Urteil geteilt, dass trotz oder gerade wegen der Digitalisierung und ihrer Folgen die Zeiten schlechter und nicht besser würden. Einen Beleg für die negativen Zukunftsaussichten liefern scheinbar die Produktivitätsfortschritte. Sie haben sich in den USA und Europa in den letzten Dekaden stetig verlangsamt.

An dieser Stelle setzt der Beitrag Henning Vöpels ein. Vöpel beschreibt, wie sich die makroökonomische Produktionsfunktion im Laufe der Zeit technologisch verändert hat und was das für die jeweilige Ökonomie und Ökonomik bedeutete. Daten würden zu einer neuen Ressource, was weitreichende Konsequenzen für die Produktionsfaktoren Kapital und Arbeit sowie letztlich für die Ordnung der neuen Digitalökonomie habe.

In Fortführung der historischen Betrachtung hinterfragt Thomas Straubhaar in seinem Beitrag, ob und gegebenenfalls inwieweit das gegenwärtige Produktivitätsparadox mit dem Strukturwandel von Industrie- zu Datenökonomien zu tun hat. Wenn bis anhin Unsichtbares sichtbar und damit erfassbar würde, dann könnte sich herausstellen, dass westliche Volkswirtschaften kein tatsächliches Produktivitäts-, sondern vielmehr ein (statistisches) Messproblem haben.

Der Beitrag Felix Roths liefert weitere Argumente dafür, dass das Produktivitätsparadox mit Strukturwandel erklärt werden kann, und gibt einen Ausblick auf die COVID-19-Krise. Er bietet zwei Hauptschlussfolgerungen. Zunächst wird der Schluss gezogen, dass ein großer Teil des Produktivitätsrätsels gelöst werden kann, indem immaterielles Kapital in die Vermögensgrenze der Volkswirtschaftlichen Gesamtrechnung einbezogen wird. Zweitens sieht der Beitrag ein Post-COVID-19-Szenario vor, das höchstwahrscheinlich zu einem deutlichen Anstieg des Arbeitsproduktivitätswachstums führen wird. Dies hängt jedoch davon ab, ob der gegenwärtig stattfindende Digitalisierungsschub mit Investitionen in die Digitalisierung sowie das notwendige ergänzende (unternehmerische und öffentliche) immaterielle Kapital flankiert wird.

Henrique Schneider stellt in seinem Beitrag ein zweites „Rätsel der Produktivität" fest: Die Einschränkungen der Wirtschaft zur Eindämmung der Pandemie im ersten Halbjahr 2020 haben zu Produktivitätserhöhungen geführt. Dieser kontraintuitive Befund könnte ein Zeichen für die Entfernung der volkswirtschaftlichen Statistik von der Realität der Daten- und Dienstleistungswirt-

schaft sein. Doch er könnte auch darauf hinweisen, wie viel Potenzial gerade in der Hybridisierung noch vorhanden ist.

Der zweite Teil des Sammelbands ist der Suche nach neuen empirischen Verfahren für die Datenwirtschaft gewidmet. Wie lassen sich mit Hilfe von Algorithmen, künstlicher Intelligenz und selbstlernenden Maschinen ökonomische Aktivitäten erkennen und messen? Wie kann eine Datenanalyse für Diagnosen und Prognosen von Verhalten(sänderungen) in Wirtschaft, Gesellschaft und Politik genutzt werden? Überall und ganz besonders in den Big-Tech-Firmen wird nach Antworten auf diese drängenden Fragen gesucht. Die Beiträge des zweiten Teils wollen diese Anstrengungen unterstützen.

Christina Heike Maaß beschreibt in ihrem Beitrag Möglichkeiten und Grenzen der „Nowcast"-Voraussagen. Dabei liegt ihr Hauptaugenmerk auf bei neuen Methoden der Big-Data-Analyse. Wie lässt sich mit einem „Measurement without Theory" alleine aus Datenströmen voraussagen, wie sich makroökonomische Größen verändern? Können also hochfrequente Echtzeitdaten ohne zugrundeliegende ökonomische Theorie nur durch Anwendung von Algorithmen zu effizienten Prognosen der Gegenwart genutzt werden? In diesem Beitrag werden der aktuelle wissenschaftliche Erkenntnisstand untersucht und das große Potenzial der Nowcast-Prognosen mit Big Data und maschinellem Lernen aufgezeigt.

In ihrem Beitrag greift Silke Sturm die Frage auf, inwieweit sich Texte als Grundlage eignen, um ökonomische und auch politische Prozesse greifbar zu machen. Texte sind die am schnellsten wachsende Datengrundlage weltweit. Die Vernetzung und ständige Verfügbarkeit über digitale Medien ermöglicht Einzelpersonen ein erhöhtes Beteiligungspotenzial. Während die Vorteile auf der Hand liegen, ist die Auswertung mit einem höheren Aufwand verbunden als bei strukturierten Datensätzen. Der Beitrag greift wichtige Auswertungsmethoden auf und stellt exemplarisch die Möglichkeiten anhand einer Analyse politischer Kommunikation in Sozialen Medien dar.

Deniz Dilan Karaman Örsal konzentriert sich in ihrem Beitrag auf die Verwendung sowohl von Textnachrichten aus den Sozialen Medien als auch von Ergebnissen von Onlinesuchanfragen. Mit ihrem Beitrag zeigt sie, dass Textdaten aus Sozialen Medien (wie zum Beispiel Twitter oder Facebook) und Ergebnisse von Onlinesuchanfragen (zum Beispiel Google Trends) zunehmend in der Forschung und in der Praxis angesetzt werden, um makroökonomische Variablen vorauszusagen. Darüber hinaus stellt sie Prognoseverfahren vor, welche auf der Grundlage von Onlinedaten aus den Sozialen Medien und Suchmaschinen Vor-

aussagen von Konsumentscheidungen erlauben. Sie verdeutlicht, dass sich mit dem Einsatz von Onlinedaten die Prognoseverfahren verbessern lassen und die Entwicklungen in den Konsumentscheidungen zeitgleich abschätzbar werden.

Im Beitrag von Niklas Scheffer, Silke Sturm und Zahurul Islam geht es um implizite Verhaltensmotive in der politischen Kommunikation. Implizite Motive spielen eine große Rolle dabei, wie unser Verhalten initiiert, aufrechterhalten und gestoppt wird. Sie orientieren, wählen und aktivieren unser Verhalten. Insbesondere im US-amerikanischen Kontext konnte bereits nachgewiesen werden, dass sich über die Auswertung impliziter Motive Krisen oder sogar kriegerische Auseinandersetzungen prognostizieren lassen. Der Beitrag stellt die Auswertung der Social-Media-Kommunikation deutscher Parteien (hinsichtlich ihrer drei Hauptmotive) durch maschinelles Lernen dar.

Alles in allem liefern die in dem vorliegenden Sammelband vereinten Beiträge einen Einstieg in eine Neuvermessung der Datenökonomie. Selbstredend ist das nicht das Ende, sondern bestenfalls der Anfang einer weiterführenden Analyse. Es gilt aufzudecken, wie tiefgreifend Big-Data-Ökonomie und Ökonomik im 21. Jahrhundert verändern (werden).

Literatur

Azkan, Can; Goecke, Henry; Spiekermann, Markus (2020): Forschungsbereiche der Datenökonomie. In: Wirtschaftsdienst, 100. Jg., H. 2, S. 124–127, DOI https://doi.org/10.1007/s10273-020-2582-x.

Brynjolfsson, Erik et al. (2019): GDP-B.: Accounting for the Value of New and Free Goods in the Digital Economy. Cambridge, Mass.: NBER Working Paper 25695, DOI https://doi.org/10.3386/w25695.

Coyle, Diane (2014): GDP. A brief but affectionate history. Princeton University Press, New Jersey.

Dewenter, Ralf und Lüth, Hendrik (2016): Big Data aus wettbewerblicher Sicht. In: Wirtschaftsdienst, 96. Jg., H. 9, S. 648–654, DOI https://doi.org/10.1007/s10273-016-2029-6.

Economist (2017): The world's most valuable resource is no longer oil, but data. Ausgabe vom 6.5.2017. Online abrufbar unter https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data, zuletzt aufgerufen am 13.4.2021.

Einav, Liran und Levin, Jonathan (2014): Economics in the age of big data. In: Science, Vol. 346, Issue 6210, Nov. 2014, S. 715. Online abrufbar unter https://science.sciencemag.org/content/346/6210/1243089.long, zuletzt aufgerufen am 13.4.2021.

Enquete-Kommission (2013): Wachstum, Wohlstand, Lebensqualität – Schlussbericht. Deutscher Bundestag, Drucksache 17/13300 vom 3.5.2013. Online abrufbar unter http://dip21.bundestag.de/dip21/btd/17/133/1713300.pdf, zuletzt aufgerufen am 13.4.2021.

Enste, Dominik H. (2019): Verluste der Unternehmen durch Schwarzarbeit. In: Wirtschaftsdienst, 99. Jg., H. 2, S. 152–154, DOI https://doi.org/10.1007/s10273-019-2411-2.

Finanzen100 (2021): Die Top 100 größten börsennotierten Unternehmen der Welt. Online abrufbar unter https://www.finanzen100.de/top100/die-grossten-borsennotierten-unternehmen-der-welt/, zuletzt aufgerufen am 13.4.2021.

Garboden, Philip M. E. (2020): Sources and Types of Big Data for Macroeconomic Forecasting. In: Peter Fuleky (Hrsg.): Macroeconomic Forecasting in the Era of Big Data. Theory and Practice. Springer International Publishing, S. 3–23, DOI https://doi.org/10.1007/978-3-030-31150-6_1.

Harding, Matthew und Hersh, Jonathan (2018): Big Data in economics (New sources of data create challenges that may require new skills). IZA World of Labor, September 2018, DOI https://doi.org/10.15185/izawol.451.

Haucap, Justus (2018): Daten als Wettbewerbsfaktor. In: Wirtschaftsdienst, 98. Jg., H. 7, S. 472–477, DOI https://doi.org/10.1007/s10273-018-2318-3.

Hübl, Lothar (2007): Wirtschaftskreislauf und gesamtwirtschaftliches Rechnungswesen. In: Vahlens Kompendium der Wirtschaftstheorie und Wirtschaftspolitik, Band 1, 9. Auflage. Verlag Franz Vahlen, München.

IMF (2019): The Economics of Data. Online abrufbar unter https://blogs.imf.org/2019/09/23/the-economics-of-data, zuletzt aufgerufen am 13.4.2021.

Konzeptwerk Neue Ökonomie (2017): Kein Wachstum ist auch (k)eine Lösung. Mythen und Behauptungen über Wirtschaftswachstum und Degrowth. Rosa-Luxemburg-Stiftung Berlin (luxemburg argumente Nr. 14). Online abrufbar unter https://www.rosalux.de/publikation/id/14854/kein-wachstum-ist-auch-keine-loesung, zuletzt aufgerufen am 13.4.2021.

Meadows, Donella H.; Meadows, Dennis L.; Randers, Jørgen; Behrens, William W. (1972): The Limits to Growth. A report for the Club of Rome's project on the predicament of mankind. Washington DC (Potomac Associates Books). Online abrufbar unter https://collections.dartmouth.edu/teitexts/meadows/diplomatic/meadows_ltg-diplomatic.html, zuletzt aufgerufen am 13.4.2021.

Ochs, Carsten; Friedewald Michael; Hess, Thomas; Lamla, Jörn (Hrsg.; 2019): Die Zukunft der Datenökonomie. Zwischen Geschäftsmodell, Kollektivgut und Verbraucherschutz. Springer Verlag, Wiesbaden, DOI https://doi.org/10.1007/978-3-658-27511-2.

Organisation for Economic Cooperation and Development (OECD; 2014): Understanding National Accounts. Paris (OECD Publishing), DOI https://doi.org/10.1787/9789264214637-en.

Plickert, Philip (2014): Sex, Drogen und Waffen für das BIP. In: Frankfurter Allgemeine Zeitung vom 11.8.2014. Online abrufbar unter https://www.faz.net/aktuell/wirtschaft/wirtschaftspolitik/illegale-aktivitaeten-sex-drogen-und-waffen-fuer-das-bip-13090550.html, zuletzt aufgerufen am 13.4.2021.

Räth, Norbert (2016): Volkswirtschaftliche Gesamtrechnungen – Reflexionen 2016. In: Wirtschaft und Statistik, Juni 2016, H. 3, S. 96–113. Online abrufbar unter https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001414/VGR.pdf, zuletzt aufgerufen am 13.4.2021.

Räth, Norbert und Braakmann, Albert (2014): Generalrevision der Volkswirtschaftlichen Gesamtrechnungen 2014 für den Zeitraum 1991 bis 2014. In: Statistisches Bundesamt, Wirtschaft und Statistik, Oktober, H. 5, S. 502–543. Online abrufbar unter https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2014/09/generalrevision-1991-2014-0912014.pdf?__blob=publicationFile, zuletzt aufgerufen am 13.4.2021.

Saetnan, Ann; Schneider, Ingrid; Green, Nicola (Hrsg.; 2018): The Politics of Big Data (Big Data, Big Brother). Routledge, Oxford und New York.

Schmelzer, Matthias (2016): The Hegemony of Growth. The OECD and the Making of the Economic Growth Paradigm. Cambridge University Press, Cambridge (UK).

Schneider, Friedrich und Boockmann, Bernhard (2020): Die Größe der Schattenwirtschaft – Methodik und Berechnungen für das Jahr 2020. Institut für Angewandte Wirtschaftsforschung, Tübingen.

Statistisches Bundesamt (2017): Bruttoinlandsprodukt 2016 für Deutschland. Begleitmaterial zur Pressekonferenz am 12. Januar 2017 in Berlin. Wiesbaden. Online abrufbar unter https://www.destatis.de/DE/Presse/Pressekonferenzen/2017/BIP2016/pressebroschuere-bip.html, zuletzt aufgerufen am 13.4.2021.

Stobbe, Alfred (1994): Volkswirtschaftliches Rechnungswesen, 8. Auflage. Springer-Verlag, Berlin/Heidelberg, DOI https://doi.org/10.1007/978-3-642-57941-7.

Stobbe, Alfred (1988): Volkswirtschaftliche Gesamtrechnung. In: Handwörterbuch der Wirtschaftswissenschaften, Band 8. Gustav Fischer / JCB Mohr, Stuttgart/Tübingen, S. 368–405.

Stone, Richard (2013): The Role of Measurement in Economics. Cambridge: Cambridge University Press, DOI https://doi.org/10.1017/CBO9781316530283.

Tagesschau (2021): Wertvollste Unternehmen. Deutsche Konzerne abgehängt. Online abrufbar unter https://www.tagesschau.de/wirtschaft/wertvollste-unternehmen-deutschlands-firmen-unter-ferner-liefen-101.html, zuletzt aufgerufen am 13.4.2021.

Taschowsky, Peter (2015): Illegale Aktivitäten in den Volkswirtschaftlichen Gesamtrechnungen. Möglichkeiten und Grenzen der Erfassung von Zigarettenschmuggel und Drogen. In: Wirtschaft und Statistik, April 2015, H. 2, S. 28–41. Online abrufbar unter https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001110/IllegaleAktivitaetenVGR_2_2015.pdf, zuletzt aufgerufen am 13.4.2021.

United Nations Development Programme (UNDP; 2020): Human Development Reports. Online abrufbar unter http://hdr.undp.org/en, zuletzt aufgerufen am 13.4.2021.

Varian, Hal R. (2014): Big Data: New Tricks for Econometrics. In: Journal of Economic Perspectives 28 (2), S. 3–28, DOI https://doi.org/10.1257/jep.28.2.3.

Wambach, Achim; Müller, Hans Christian (2018): Digitaler Wohlstand für alle. Ein Update der Sozialen Marktwirtschaft ist möglich. Campus Verlag, Frankfurt am Main/New York: .

Teil 1: Das Produktivitätsparadox der Datenökonomie

Die digitale Revolution: Der große Übergang in die Datenökonomie

Henning Vöpel

Die sogenannte digitale Revolution beginnt mit einem technologischen Sprung: Erstmals in der Geschichte der Menschheit ist es möglich, massenhaft Daten – „Big Data" – zu speichern, zu verarbeiten und miteinander zu verknüpfen, und zwar in Echtzeit und mobil an fast jedem Ort der Welt. Mehr als vier Milliarden Menschen sind bereits heute per Smartphone erreichbar. Die Welt ist weder Scheibe noch Kugel: Sie ist technologisch und ökonomisch zu einer Plattform geworden. Nicht überraschend also, dass auch die Globalisierung eine neue, nämlich digitale Qualität der Vernetzung erfährt.[1]

Dieser zunächst harmlos anmutende technologische Sprung löst eine geradezu epochale ökonomische und auch kulturelle Revolution aus. Denn Daten können nun genutzt werden, um Menschen, Maschinen und Infrastrukturen miteinander zu verbinden – alles wird „smart". Ökonomisch betrachtet werden Daten dadurch zu einer Ressource, in der hohe Wertschöpfungspotenziale liegen und die nun systematisch gehoben werden können. Plattformen und Algorithmen sind die Vehikel dafür. Was noch harmlos klingt, verändert jedoch nicht weniger als die Ordnung der Wirtschaft.[2] Der folgende Beitrag ist eine grundsätzliche Einführung in die Datenökonomik aus produktions- und allokationstheoretischer Perspektive.

---

[1] Vgl. allgemein zur Digitalökonomie zum Beispiel Goldfarb/Tucker (2019).

[2] Vgl. Vöpel (2018).

Die technologische Evolution der Ökonomie: Der Übergang von der Industriewirtschaft in die Datenökonomie

Die Digitalisierung kann als zweite große technologische Revolution in der Wirtschaftsgeschichte verstanden werden. Die erste große Revolution war der Übergang von der Agrarwirtschaft in die Industriewirtschaft. Die zweite ist nun der Übergang von der Industriewirtschaft in die Digitalwirtschaft. Sehr stark stilisiert lässt sich die Evolution der Wirtschaft anhand einer makroökonomischen Produktionsfunktion darstellen (vgl. Abbildung 1).

In der Agrarwirtschaft war der wesentliche spezifische Faktor der „Boden". Eigentum an Boden bedeutete die Möglichkeit, die ökonomische Knappheitsrente abzuschöpfen, denn der Boden war nicht vermehrbar, sodass Produktivitätsfortschritte nicht unwesentlich diesem zufielen. Entsprechend waren nicht nur die ökonomischen, sondern auch die gesellschaftlichen und politischen Verhältnisse durch den Grundbesitz bestimmt.

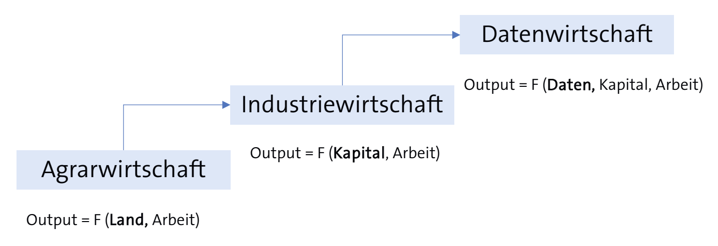


Abb. 1: Die Evolution der Wirtschaft in der Produktionsfunktion
Quelle: Schneider/Vöpel/Weis (2018).

Mit der Industrialisierung löste – etwas überspitzt formuliert – der Fabrikbesitzer den Großgrundbesitzer als Kapitalisten ab. Kapital wurde zum bestimmenden, weil relativ knappen Produktionsfaktor, der Faktor Arbeit dagegen, der vom Agrar- in den Industriesektor wechselte, musste anfangs – noch ungelernt und daher im Wettbewerb mit einer ungeheuren „Reservearmee" aus Arbeitern stehend – zu sehr geringen Löhnen arbeiten. Die Arbeitswertlehre von Karl Marx sah in Kapital sogar „geronnene Arbeit"; Kapital könne nur deshalb einen Zins erwirtschaften, weil zuvor Arbeit ausgebeutet worden sei.[3] Die politische und gesellschaftliche Entwicklung der Wirtschaft war jedenfalls immer stark dadurch geprägt, wer Eigentum an der ökonomisch knappen Ressource hielt.

---

[3] Vgl. hierzu Marx (1867), Sraffa (1975), Dow (1993).

Das gilt auch für den zweiten großen technologischen Übergang, jenen von der Industrieökonomie in die Datenökonomie. Auch dieser kommt einer Revolution gleich, die nicht nur eine technologische, sondern zugleich eine ökonomische, gesellschaftliche und kulturelle ist: Daten lösen das physische Kapital als den spezifischen Produktionsfaktor ab. Die makroökonomische Produktionsfunktion der digitalen Datenökonomie weist somit drei Produktionsfaktoren auf: Arbeit, Kapital und eben Daten. Der neue Kapitalist hat nunmehr Eigentum an Daten. Durch Daten als neuem Produktionsfaktor verändern sich zunächst die produktionstheoretischen Beziehungen zwischen den Produktionsfaktoren.

Daten versus Kapital

Daten als Produktionsfaktor bedeuten, dass sich die Produktion und die Produkte dematerialisieren, sie begründen die intangible Ökonomie.[4] Zwar werden immer noch Kapital und Arbeit benötigt, jedoch liegen in den Daten die nunmehr spezifischen, nicht durch andere Produktionsfaktoren substituierbaren Informationen, weshalb den Daten – wie im klassischen „Kapitalismus“ dem Kapital – die ökonomische Rente zufällt. Dieser Vorgang ist bereits deutlich zu beobachten: Große kapitalintensive Unternehmen werden durch datengetriebene Geschäftsmodelle quasi aus dem Stand disruptiert, das physische Kapital wird entwertet. Folgerichtig sind heute Plattform-unternehmen wie Amazon oder Google die am stärksten marktkapitalisierten Unternehmen der Welt und haben längst die großen Energie- oder Stahlunternehmen abgelöst.

Daten versus Arbeit

Auch das Verhältnis zwischen Daten und Arbeit verändert sich. Musste der „Arbeiter“ im Kapitalismus noch vor einer Ausbeutung seiner Arbeit durch das Kapital geschützt werden, indem sich klassische Arbeitergewerkschaften bildeten, so muss womöglich im Datenkapitalismus der Konsument vor einem Missbrauch seiner Daten geschützt werden, da er sonst den sich aufgrund der Netzwerkeffekte der Daten monopolisierenden Plattformen hilflos ausgeliefert ist. Darüber hinaus besteht zwischen Daten und Arbeit eine produktionstheoretische Substitutionsbeziehung. Der Einsatz von Daten und künstlicher Intelli-

---

[4] Vgl. Haskel/Westlake (2017).

genz kann insbesondere kognitive Routinearbeit des Menschen ersetzen, wodurch sich die Entlohnung von jahrelanger Ausbildung erheblich verringern könnte. Es besteht daher vielfach die Befürchtung, Digitalisierung könne zumindest übergangsweise zu technologischer Massenarbeitslosigkeit führen. Diese Befürchtungen sind aus historischer Perspektive unbegründet, jedoch wird die technologische Entwicklung sich womöglich so stark beschleunigen, dass auf den Arbeitsmarkt, das Bildungssystem und den Sozialstaat große Herausforderungen zukommen werden.

Digitalökonomik: Daten, Plattformen und künstliche Intelligenz

Daten gehen nicht allein als Produktionsfaktor in die makroökonomische Produktionsfunktion, sondern umfassend in den volkswirtschaftlichen Produktionsprozess ein, der allgemein als Umwandlung von Ressourcen in Güter und Dienstleistungen verstanden werden kann (vgl. Abbildung 2).



Abb. 2: Daten im Produktions- und Umwandlungsprozess
Quelle: eigene Darstellung.

Daten spielen daher auch als Outputgröße eine wichtige Rolle. Sie sind Teil neuer, individualisierter, das heißt kundenzentrierter und präferenzgerechter Dienstleistungen, sie machen selbst aus physischen Gütern wie Autos oder Wohnungen Services, indem sie deren Nutzung verstetigen und so die Ökonomie nicht zuletzt zu einer Sharing Economy machen.[5] Physische Assets lassen sich durch die Echtzeitnutzung der Daten weitaus effizienter nutzen, indem zum Beispiel Autos durch die Koordinierung und Synchronisierung von Abläufen permanent nutzbar gemacht werden können sowie Transaktions- und Suchkosten auf nahe null sinken.

Mit den Daten als Input- und Outputgröße gewinnen Algorithmen und künstliche Intelligenz sowie Plattformen als operativ-organisatorisches Vehi-

[5] Vgl. zum Beispiel Rifkin (2014).

kel der systematischen Nutzung von Daten an Bedeutung. Daten und künstliche Intelligenz sind in diesem Sinne komplementäre Produktionsfaktoren. Die meisten Informationstechnologien können nur in Kombination mit firmenspezifischem Kapital Wertschöpfung generieren; letzteres kann als digitales Kapital bezeichnet werden.[6] So überrascht es nicht, dass die großen Plattformen nicht nur branchen- und nutzerübergreifend enorme Mengen an Daten sammeln, sondern zugleich die größten Investoren in künstliche Intelligenz sind. Die Informationen in Daten veralten schnell, sodass deren Nutzung an die unmittelbare Auswertung gekoppelt ist. Das wiederum ist nur möglich, wenn Algorithmen in der Lage sind, ungeheure Datenmengen in Nanosekundenschnelle auszuwerten. In der Produktion führen Daten und Algorithmen dazu, dass die Skalenvorteile der Massenproduktion geringer werden und geringere Losgrößen nun weit kostengünstiger produziert werden können.

Daten weisen als Ressource besondere ökonomische Eigenschaften auf. Gelegentlich wird von Daten als dem neuen Gold oder dem neuen Öl der Datenwirtschaft gesprochen. Streng genommen ist dies falsch, denn Daten sind anders als Öl wieder-, gleichzeitig und mehrfach verwertbar, es besteht also „Nichtrivalität" in der Nutzung. Darüber hinaus können Nutzer und Mitbewerber von der Nutzung der Daten ausgeschlossen werden. Es bilden sich daher private Datenunternehmen, die Daten als private Ressourcen nutzen und damit private Güter produzieren. Es sind jedoch oftmals Kollektivgüter, die nur dadurch entstehen können, dass sich mehrere Nutzer an ihrer Erstellung beteiligen und so Netzwerkeffekte internalisieren. Eine Mobilitäts- oder Hotelzimmer-Plattform ist umso effizienter, je mehr Nutzer sich auf ihr versammeln. Aufgrund dieser Netzwerkeffekte entstehen Skalierungsvorteile datenbasierter Geschäftsmodelle. Diese werden auf einer Plattform effizient verarbeitet, denn dort können sie konzentriert und optimal miteinander verknüpft werden. Plattformen sind daher der „Industriekonzern" der Digitalökonomie.

Phänomene und Folgen der Datenökonomie

Die beschriebenen Eigenschaften von Daten als Ressource haben weitreichende ökonomische und ordnungspolitische Folgen. Oberflächlich haben sich bereits offensichtliche Phänomene gebildet. Nachstehend sind die wichtigsten zusammengefasst und beschrieben:

---

[6] Vgl. Brynjolfsson et al. (2020).

- Es bilden sich Plattformen (Beispiele: Facebook, Amazon), da auf Plattformen Daten effizient erfasst und miteinander zu neuen datenbasierten Dienstleistungen verknüpft werden können. Plattformen werden durch die Netzwerkeffekte der Daten in der Tendenz immer größer, weil sie für Nutzer umso attraktiver sind, je größer sie sind, vor allem sogenannter zweiseitiger Märkte.[7]
- Durch Plattformen werden Intermediäre eliminiert, da Produzenten und Konsumenten direkt, das heißt über die Plattformen miteinander in Kontakt treten.[8]
- Datenbasierte Plattformen ziehen Margen von den traditionellen Produzenten und Händlern ab, da die Information über die Konsumenten sowie deren Präferenzen und Verhaltensmuster nun bei den Plattformen liegt.[9]
- Physische Güter werden auf Plattformen durch die Anreicherung von Daten zu digitalen Dienstleistungen (… as a service). Intangible Assets gewinnen in der Digitalökonomie dadurch massiv an Bedeutung.
- Das physische Kapital entwertet sich durch diese Dematerialisierung zugunsten der Plattformen (Beispiele: Uber, Airbnb).

Jenseits dieser eher oberflächlichen Erscheinungen verändert sich in der Digitalökonomie die Ordnung der Wirtschaft und mit ihr die Art und Weise, wie Wirtschaft organisiert ist. In der Industrialisierung haben sich vertikale Wertschöpfungsketten gebildet. Die großen Treiber waren Arbeitsteilung und Spezialisierung. Das klassische Industrieunternehmen war daher ein Teil einer komplexen Wertschöpfungskette, welches die Vorteile der Spezialisierung realisierte und über transaktionale Beziehungen zu Zulieferern und Kunden die Wertschöpfungskette organisierte. Ronald Coase hat diese Form des Unternehmens in seiner Theorie der Unternehmung eingehend beschrieben.[10] Arbeitskräfte haben sich ebenfalls spezialisiert, da die Entlohnung mit dem Spezialisierungsgrad der Arbeit und entsprechend höherer Arbeitsproduktivität steigt. Das Prinzip vertikaler Organisation industrieller Produktion ist stark stilisiert in Abbildung 3 dargestellt.

---

[7] Rysman (2009).

[8] Vgl. Palmatier/Moorman/Lee (2019).

[9] Vgl. Brynjolfsson/Hu/Rahman (2009) und Brynjolfsson/Oh (2012).

[10] Vgl. Coase (1937).

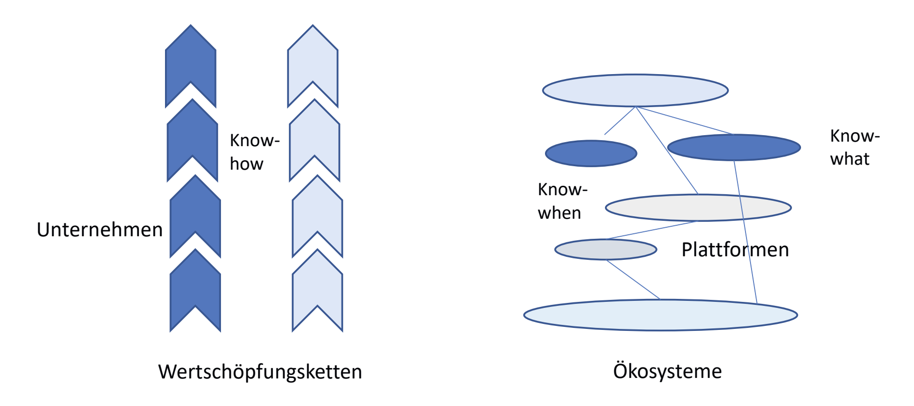


Abb. 3: Vertikale Wertschöpfungsketten versus digitale Wertschöpfungsnetzwerke
Quelle: Schneider et al. (2018).

Die Wertschöpfung in der Digitalökonomie ist somit nicht mehr vertikal organisiert, sondern wird sich gewissermaßen „diagonalisieren", das heißt, Branchengrenzen lösen sich auf und neue Geschäftsmodelle entstehen „zwischen" den klassischen Branchen. Wesentlich hierfür sind wiederum Daten, die zwischen den Branchen und Unternehmen Schnittstellen reduzieren und Innovationspotenziale für neue Prozesse und Produkte schaffen. Die hochspezialisierten, arbeitsteilig organisierten Unternehmen werden in der Digitalökonomie zu hybriden, offenen Strukturen, die in der Lage sind, agil und kollaborativ Wertschöpfungspotenziale zu erkennen und zu heben. Damit ist indes ein tiefer Kultur- und Führungswandel in den tradierten Unternehmen verbunden. Digitale Wertschöpfung wird in Wertschöpfungsnetzwerken stattfinden, in denen die Erfassung und die Zuordnung von Wertschöpfung kaum mehr möglich sein werden. Daran knüpfen sich grundlegende rechtliche und steuersystematische Fragen, die hier erwähnt, aber nicht näher erläutert werden sollen, da sie über den Zweck dieses Beitrags weit hinausgehen.

Darüber hinaus werden einige „klassische" Axiome und Theoreme der Ökonomik in Frage gestellt. Abnehmende Grenzerträge oder das Prinzip komparativer Kostenvorteile und daraus abgeleitete Schlussfolgerungen für Spezialisierung, Arbeitsteilung und Handel sind davon betroffen. Insbesondere künstliche Intelligenz kann dazu führen, dass Wissen und Fähigkeiten sich replizieren und dadurch überall auf der Welt verfügbar gemacht werden können. Es liegt daher die Vermutung nahe, dass die Digitalökonomik die Art und Weise, wie die

globale Wirtschaft sich organisieren wird, radikal verändern und sogar fundamentale Ordnungsprinzipien in Frage stellen wird.

Ordnungsfragen der Datenökonomie

Die Eigenschaften von Daten als Ressource und deren Nutzung durch Plattformen und Algorithmen werfen grundlegende Ordnungsfragen auf, die heute im Wesentlichen unbeantwortet und nicht befriedigend gelöst sind. Die wettbewerbsökonomische oder zumindest ordnungspolitische Bewertung ist zunächst ambivalent.[11] Zum einen ergeben sich durch die Plattformen enorme Effizienzpotenziale in der Koordinierung und Individualisierung von Tausch- und Produktionsprozessen, zum anderen führen die Skalen- und Netzwerkeffekte der Daten zu einer Monopolisierung in der Datenökonomie, wodurch eine Re-Regulierung und eine Novellierung des Wettbewerbsrechts notwendig werden wie zugleich die Etablierung eines umfassenden Datenrechts, das Aspekte des Datenschutzes, der Datensicherheit und der informationellen Selbstbestimmung beinhaltet. Diese Fragen können hier nicht beantwortet, die Probleme nicht gelöst werden. Einige Bereiche können jedoch skizziert und angerissen werden.

Die wichtigste Frage im digitalen Zeitalter stellt sich in Bezug auf die Ökonomisierung von Daten: Wem gehören die Daten und welche Daten sollen wann, von wem und zu welchem Zweck genutzt werden können? Hier gilt es, den Konsumenten die Hoheit über die Daten zurückzugeben. Dies ist kein leichtes Unterfangen, denn die Plattformen sind bereits so groß, dass sie Lock-in-Effekte erzeugt haben, die es für die Konsumenten teuer machen, den kundenzentrierten Angeboten fernzubleiben. Generell lässt sich sagen, dass in Demokratien die Nutzung von Daten wohl weitaus restriktiver gehandhabt werden wird als in Autokratien. Grundsätzlich aber werden Gesellschaften und vor allem Individuen deutlich gläserner werden. Die Entscheidung darüber, wie gläsern, sollte jedoch weitgehend bei den Individuen selbst liegen.

Die zweite Frage ist unmittelbar damit verknüpft: Wie lassen sich Datenschutz und Datensicherheit gewährleisten? Während große Datenmengen aufgrund ihrer Netzwerkeffekte enorme Effizienzpotenziale offerieren, nehmen zugleich damit systemische Risiken der Datensicherheit und des Datenschutzes zu. Im Jahr 2020 kam es zu einem „Datenleck" in einer psychiatrischen Anstalt,

---

[11] Vgl. allgemein zum Beispiel Varian (2005).

wodurch persönliche, sehr sensible Daten von Hunderten von Patienten öffentlich wurden. Der ehemalige Vorsitzende des Deutschen Ethikrats, Peter Dabrock, bezeichnete diesen Vorgang als das „Tschernobyl der Digitalisierung“[12]. Die systemischen Risiken von Cyberattacken oder Serverausfällen sollten durch eine entsprechende Dateninfrastruktur kontrolliert werden, denn diese reduziert die variablen Kosten für die Anwender und könnte so die Digitalisierung und deren Anwendungen wesentlich beschleunigen.

Die dritte Frage bezieht sich auf das erste Vehikel der Datenökonomie, die Algorithmen und die künstliche Intelligenz. Diesbezüglich stellt sich die Frage: Wie lässt sich Transparenz über die Algorithmen herstellen? Diese Frage ist deshalb so wichtig, weil Algorithmen persönliche Daten sammeln, filtern und auswerten und so Konsumentscheidungen beeinflussen und manipulieren können. Transparenz über die Art und den Umfang des Einsatzes von Algorithmen herzustellen ist deshalb ein wichtiger Aspekt der digitalen Regulierung. Besteht diese Transparenz nicht, drohen Plattformen ihre Nutzer zu kontrollieren, vielleicht sogar zu manipulieren und zu diskriminieren. Shushana Zuboff nennt dies einen entmündigenden Überwachungskapitalismus.[13]

Die vierte Frage schließlich betrifft die Plattformen als Vehikel der datenbasierten Wertschöpfung selbst, denn diese haben die Tendenz, sich zu monopolisieren. Zwar waren auch klassische Industrieunternehmen oft durch Skalierungsvorteile gekennzeichnet, diese nahmen jedoch irgendwann ab, während die Netzwerkeffekte der Daten immer weiter zunehmen, je größer Plattformen werden. Die vielleicht entscheidende Frage der Digitalisierung lautet: Wie lassen sich Plattformen sinnvoll regulieren, sodass deren Effizienz weiterhin erhalten bleibt, ihre Marktmacht jedoch begrenzt wird? Bislang gab es vonseiten der Regulierungsbehörden überwiegend Vorschläge zur Begrenzung der Marktmacht, zur Sanktionierung des Missbrauchs von Marktmacht oder zu einer verschärften Fusionskontrolle. Auch eine Zerschlagung von zu großen und marktmächtigen Plattformen wird immer wieder ins Spiel gebracht. Selbst große Plattformen sind jedoch nicht zwingend weniger innovativ und kundenfreundlich als Unternehmen auf wettbewerblichen Märkten, solange die Märkte bestreitbar bleiben („contestable markets“), die Markteintrittskosten also gering sind, was bei kapitallosen Geschäftsmodellen oft der Fall ist.

---

[12] Vgl. Dabrock (2020).

[13] Vgl. Zuboff (2019).

Zudem findet der digitale Wettbewerb weniger im Güterraum statt als vielmehr im Datenraum. So konkurrieren heute etwa Google, Amazon oder Facebook um Paymentlösungen, obgleich jede Plattform für sich eine eigene Marktführerschaft innehat. Die Wettbewerbsintensität zwischen Plattformen kann erheblich erhöht werden, wenn diese interoperational sind, die Regulierung es also für Nutzer leichter macht, zwischen unterschiedlichen Plattformen zu wechseln, sodass Lock-in-Effekte vermieden werden. Öffentlich betriebene und bereitgestellte Plattformen – im Sinne einer oben skizzierten Plattform-Infrastruktur – sind eher keine Lösung, eher schon genossenschaftliche Formen der Datennutzung, da Daten als eine Art Kollektivressource verstanden werden können.

Schlussfolgerungen

Die digitale Revolution wird zu Recht als solche bezeichnet. Denn es handelt sich nicht um lediglich inkrementelle Veränderungen. Die Datenökonomie stellt vielmehr grundlegende Ordnungsfragen, die heute bei weitem nicht gelöst sind. Nicht allein die disruptiven Veränderungen der Geschäftsmodelle charakterisieren den Übergang von der Industrieökonomie in die Datenökonomie. Gesellschaftliche, politische, (wettbewerbs- und daten)rechtliche, regulatorische und kulturelle Veränderungen machen diesen Übergang zu einer epochalen Aufgabe, gravierende Veränderungen auf dem Arbeitsmarkt sowie in der Bildung und im Sozialstaat eingeschlossen.

Literatur

Brynjolfsson, Eric; Hu, Y.; Rahman, M. S. (2009): Battle of the retail channels. How product selection and geography drive cross-channel competition. In: Management Science 55 (11), S. 1755–1765.

Brynjolfsson, Eric; Oh, J. (2012): The Attention Economy: Measuring the Value of Free Digital Services on the Internet. In: ICIS Proceedings 2012, International Conference on Information Systems der Association for Information Systems Research.

Brynjolfsson, Eric; Hitt, Lorin; Rock, Daniel; Tambe, Prasanna (2020): Digital Capital and Superstar Firms, NBER Working Paper Nr. 28285.

Coase, Ronald (1937), The Nature of the Firm. In: Economica, Vol. 4 (16,) S. 386–405.

Dabrock, Peter [just_ethics] (2020): Nachricht könnte zum #Tschernobyl der Digitalwelt mutieren. … [Tweet]. Twitter. https://twitter.com/just_ethics/status/1321360310155071488, 28. Oktober.

Dow, Gregory K. (1993): Why Capital Hires Labor. A Bargaining Perspective. In: American Economic Review, Vol. 38 (1), S. 118–134.

Goldfarb, Avi; Tucker, Catherine (2019): Digital Economics. In: Journal of Economic Literature, 57(1), S. 3–43.

Haskel, Jonathan; Westlake, Stian (2017): Capitalism without Capital. The Rise of the Intangible Economy. Princeton University Press, Princeton.

Marx, Karl (1867): Das Kapital. Eine Kritik der politischen Ökonomie. Otto-Meissner-Verlag, Hamburg.

Palmatier, Robert W.; Moorman, Christine, Lee, Ju-Yeon (2019): Handbook of Customer Centricity. Strategies for Building a Customer-centric Organization, Edward Elgar, Cheltenham/Northampton.

Rifkin, Jeremy (2014): Die Null-Grenzkosten-Gesellschaft, Campus Verlag, Frankfurt am Main / New York.

Rysman, Marc (2009): The Economics of two-sided Markets. In: Journal of Economic Perspectives, Vol. 23 (3), S. 125–143.

Schneider, Johannes; Vöpel, Henning; Weis, Martin (2018): Künstliche Intelligenz und die Neuordnung der Wirtschaft, Wien.

Sraffa, Piero (1975): Production of Commodities by Means of Commodities. Cambridge University Press, Cambridge (UK).

Varian, Hal R. (2005). Copying and copyright. In: Journal of Economic Perspectives 19 (2), S. 121–138.

Vöpel, Henning (2018), Wie Künstliche Intelligenz die Ordnung der Wirtschaft revolutioniert. In: Wirtschaftsdienst, 98 (11), S. 828–830. Online abrufbar unter https://www.wirtschaftsdienst.eu/inhalt/jahr/2018/heft/11/beitrag/wie-kuenstliche-intelligenz-die-ordnung-der-wirtschaft-revolutioniert.html.

Zuboff, Shoshana (2019): The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power. Public Affairs, New York.

Der Rückgang des Produktivitätsfortschritts: Worum geht es?

Thomas Straubhaar

Einleitung

„Technologischer Fortschritt ist der Motor für langfristiges Wirtschaftswachstum." So kurz und knapp formulierte das Stockholmer Nobelpreiskomitee den theoretischen Stellenwert neuer Technologien, als es 1987 Robert Solow für seine wissenschaftlichen Arbeiten über ökonomische Wachstumstheorien auszeichnete.[1] Und weiter: „Wenn nach Solows Modell ein kontinuierlicher technologischer Fortschritt angenommen werden kann, wird das Wachstum der Realeinkommen ausschließlich durch den technologischen Fortschritt bestimmt." Also müsste im Zeitalter der Computer, der zunehmend leistungsfähigeren Automaten, der Roboter, der künstlichen Intelligenz, der autonomen und damit wirklich auto-mobilen Fahrzeuge sowie der selbstlernenden Maschinen ein Wachstumsschub die hochentwickelten Volkswirtschaften immer schneller voranbringen.

Die empirische Realität jedoch scheint die theoretischen Erwartungen zu widerlegen. Das erkannte gerade der für seine Wachstumsmodelle geehrte Robert Solow: „Was alle für eine technologische Revolution halten – eine drastische Verbesserung in unserem produktiven Leben – wurde überall, einschließlich Japan, von einer Verlangsamung des Produktivitätswachstums begleitet, nicht von einer Steigerung. Überall sind Computer zu sehen, nur nicht in den Statistiken der Produktivitätsmessung."[2] Stand bei Robert Solows Feststellung das Digitalisierungszeitalter noch am Anfang – und bezog er sich neben der Telekommunikation insbesondere auf die programmierbare Automation – bestä-

---

[1] Royal Swedish Academy of Sciences (1987; übersetzt von T. S.).

[2] Solow (1987), S. 36.

tigt sich die Diskrepanz zwischen Theorie und Praxis auch im Übergang zu einer Datenökonomie: „We see the influence of the information age everywhere, except in the GDP statistics", so Brynjolfsson und Saunders.[3]

Der Widerspruch zwischen theoretischer Erwartung und tatsächlicher Erfahrung wird als Produktivitätsparadox bezeichnet.[4] Damit soll beschrieben werden, dass sich in den letzten Dekaden das Wachstum sowohl der Arbeits- als auch der Totalen Faktorproduktivität (TFP) trotz der fortlaufenden technologischen Innovationen trendmäßig verlangsamt hat.[5] Durch neue Technologien wurde das Tempo des technischen Fortschritts offenbar in der Realität nicht beschleunigt, sondern gebremst.

Das Produktivitätsparadox ist keine Thematisierung einer abstrakten akademischen Bagatelle. Es liefert Stoff für eine höchst relevante wirtschaftspolitische Kontroverse. Denn Konsequenz des negativen Einflusses auf die Produktivitätsentwicklung wäre es, Innovationen weit kritischer zu beurteilen, als das gemeinhin der Fall ist. Zeiten mit raschem technologischem Fortschritt wären dann nicht notwendigerweise Jahre mit hohem ökonomischem Wachstum, sondern könnten genauso mit Stagnation einhergehen. Genau an dieser Stelle steht die aktuelle Diskussion seit Beginn der 2000er-Jahre – in einer Phase, in der die Statistiken trotz Roboterisierung und Automation oder künstlicher Intelligenz und selbstlernenden Maschinen eine Verlangsamung der Produktivitätsfortschritte wiedergeben.

Die folgende Analyse will zunächst das Produktivitätsparadox als empirisches Phänomen darstellen. Danach werden Erklärungen dafür gesucht, warum die einschlägigen Statistiken in der Tat verlangsamte und nicht beschleunigte Produktivitätsfortschritte widergeben. Ein weiterer Abschnitt benennt ein paar

---

[3] Brynjolfsson und Saunders (2009).

[4] Askenazy et al. (2016).

[5] Die gesamtwirtschaftliche Arbeitsproduktivität ergibt sich aus dem Verhältnis von Bruttoinlandsprodukt (Bruttowertschöpfung) und dem Einsatz des Faktors Arbeit (gemessen entweder durch die Anzahl der Erwerbstätigen oder der insgesamt geleisteten Arbeitsstunden [(Arbeitsvolumen)]. Die Totale Faktorproduktivität setzt die Bruttowertschöpfung in Relation zu allen eingesetzten Produktionsfaktoren. Das TFP-Wachstum ergibt sich als Restgröße: „Das Wachstum der TFP wird gemessen als die Veränderung des BIP, die nicht durch Veränderungen des Arbeits- und Kapitaleinsatzes zu erklären ist. Eine Zunahme der TFP kann auf technischen Fortschritt oder Verbesserungen der Allokation von Produktionsfaktoren in der Volkswirtschaft zurückgeführt werden." (Jones [2016], zitiert nach Sachverständigenrat [2020], S. 287). Letztlich geht es also um einen technologischen Fortschritt infolge von Prozess- und Produktinnovationen – der ein effektiveres Zusammenspiel der Produktionsfaktoren (in der Regel Arbeit und Kapital) ermöglicht (vgl. dazu insbesondere Sachverständigenrat [2015], S. 284–285).

herkömmliche Erklärungen für das Produktivitätsparadox und argumentiert, dass die „säkulare Stagnation" wenig mit fehlendem Fortschritt zu tun habe, sondern Folge des strukturellen Übergangs von einer Industrie- erst zu einer Dienstleistungs- und danach zu einer Datenwirtschaft sei.

Der darauffolgende Abschnitt postuliert, dass westliche Volkswirtschaften weniger ein Produktivitäts- als vielmehr ein Messproblem haben. Es war immer schon ein grundsätzliches Problem der Wirtschaftsstatistik, dass sie zunächst nicht in der Lage ist, die Folgeeffekte von Basisinnovationen abzubilden. Wie auch?[6] Für viele der von Raum und Material losgelösten Neuerungen der Digitalisierung fehlen schlicht (noch) die gesamtwirtschaftlichen statistischen Maßzahlen.

Der letzte Abschnitt zieht ein Fazit. Er verweist auf die Notwendigkeit, bis anhin „Unmessbares" messbar zu machen. Es bedarf einer fundamentalen Anpassung der makroökonomischen Statistik. Ansonsten droht der Wirtschaftspolitik zunehmend ein Stochern im Nebel der Unschärfe und Ungenauigkeit.

Empirie: Der Produktivitätsfortschritt wird langsamer

Der empirische Befund ist eindeutig: „Für Deutschland zeigt sich langfristig ein deutlicher Rückgang des Wachstums der Totalen Faktorproduktivität. Hauptsächlich spiegelt sich darin der verlangsamte technologische Fortschritt wider."[7] Eine mehr oder weniger ähnliche Verlangsamung der TFP zeigt sich in den meisten ökonomisch höher entwickelten Volkswirtschaften.[8]

Wie Abbildung 1 veranschaulicht, lag das durchschnittliche jährliche Wachstum der TFP in den 1970er-Jahren noch bei über zwei Prozent. In den 1980er-Jahren erreichte es 1,5 %, in den 1990er-Jahren ein Prozent. Seit der Jahrhundertwende verharrt es bei etwas mehr als 0,5 %, was auch der Prognose für den Zeitraum bis 2024 entspricht.[9]

---

[6] Vgl. dazu ausführlich und kompetent Grömling (2016). Amtliche Statistiken erheben Daten zwangsläufig auf veralteten Grundlagen. Es kann nur Bekanntes gemessen werden, wofür es „Klassifikationen" gibt. Das Neue ist so unbekannt, dass es – mindestens zu Beginn – nicht mit gängigen Konzepten erfassbar ist.

[7] Sachverständigenrat (2019), S. 95.

[8] Vgl. dazu die einzelnen Beiträge in Askenazy et al. (2016) und für die USA im Speziellen Gordon (2016).

[9] Sachverständigenrat (2019), S. 95.

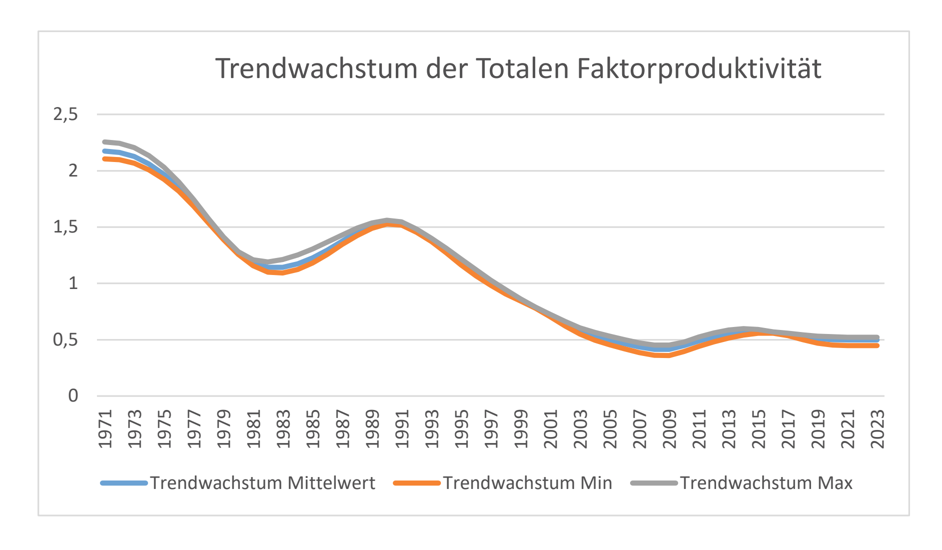


Abb. 1: Trendwachstum der Totalen Faktorproduktivität (Deutschland, 1971 bis 2024)
Quelle: Sachverständigenrat (2019), S. 95.

Kein Wunder ist es, dass angesichts der stetigen Verlangsamung des Wachstums der TFP die alte Hypothese einer „säkularen Stagnation" neue Nahrung erhält. Denn bereits vor einem Jahrhundert führte eine lange Phase schwachen ökonomischen Wachstums mit entsprechend negativen Rückwirkungen auf den durchschnittlichen Lebensstandard zu besorgtem Interesse. „Was können wir vernünftigerweise erwarten, auf welchem Niveau der Wohlstand in hundert Jahren liegen wird? Was sind die wirtschaftlichen Möglichkeiten für unsere Enkelkinder?", fragte John Maynard Keynes 1930 in Zeiten großer ökonomischer wie politischer Not und grassierender Zukunftsängste.[10]

In seinem Ausblick auf die wirtschaftlichen Chancen und Risiken für nachfolgende Generationen sah sich Keynes – angesichts der damaligen Krise nicht wirklich überraschend – einer skeptischen Perspektive seiner Zeitgenossen gegenüber: „Wir leiden gerade jetzt unter einer bösartigen Attacke des ökonomischen Pessimismus. Überall ist zu hören, dass die Epoche des enormen

[10] Keynes (1930) fasst in seinem 1930 veröffentlichten Essay „Economic Possibilities for our Grandchildren" eine Vorlesungsreihe zusammen, in der er auch seine Erkenntnisse zur Großen Depression der Zwischenkriegszeit verarbeitete.

wirtschaftlichen Fortschritts, der das vergangene Jahrhundert kennzeichnete, vorbei sei; dass die rasche Verbesserung des Lebensstandards sich fortan verlangsamen werde; dass in dem vor uns liegenden Jahrzehnt ein Rückgang des Wohlstands wahrscheinlicher sei als eine Verbesserung."[11] Führende Intellektuelle sagten der Weltwirtschaft in den 1930er-Jahren eine lange anhaltende „säkulare" Stagnation voraus.[12] Der Menschheit stehe für Dekaden ein ökonomisches Nullwachstum bevor.

Die Empirie hat den Pessimismus von damals als klare Fehlprognose entlarvt. Der Wohlstand der Massen hat nicht ab-, sondern zugenommen – und zwar weltweit. Mehr Menschen als jemals zuvor leben heutzutage besser als ihre Eltern und Vorfahren. Dennoch erscheint es, als hätte Keynes vor fast 100 Jahren nicht allein den Zukunftspessimismus von damals, sondern bereits die Ängste der Gesellschaft der Gegenwart im Auge gehabt. Denn schon lange vor dem Schrecken, den die Coronapandemie Anfang der 2020er-Jahre verbreitete, kehrten das Produktivitätsparadox und mit ihm die Sorge vor einer „säkularen Stagnation" ins Bewusstsein der Makroökonomik zurück.

So sah Lawrence Summers, einst Finanzminister unter US-Präsident Bill Clinton und später unter Barack Obama Direktor des National Economic Council, die Welt nach der Finanzmarktkrise Ende der Nullerjahre erneut vor einer langen Phase schwachen ökonomischen Wachstums mit entsprechend negativen Rückwirkungen auf den durchschnittlichen Lebensstandard.[13] Viele andere Ökonomen teilen das Urteil, dass trotz oder gerade wegen der Digitalisierung und deren Folgen der Produktivitätsfortschritt sich in den vergangenen Dekaden Schritt für Schritt verlangsamt habe und sich weiter abschwächen werde.[14] Hierfür wird eine Reihe von Ursachen verantwortlich gemacht, auf die nun näher eingegangen wird.[15]

---

[11] Keynes (1930), erster Abschnitt. Übersetzung durch T. S.

[12] Der Begriff der „säkularen Stagnation" wurde ursprünglich von dem Harvard-Ökonomen Alvin Hansen in den 1930er-Jahren geprägt (Hansen 1938) und dann von John Maynard Keynes aufgegriffen, der sich Sorgen machte, dass das durch eine – auch bereits damals – schrumpfende Bevölkerung verursachte Ungleichgewicht zwischen Sparen und Investieren zu einem längerfristigen Wachstumsschwund führen könnte.

[13] Summers (2014, 2016).

[14] Gordon (2014) ist dabei einer der führenden Vertreter der Stagnationsthese.

[15] In der Studie der Bertelsmann Stiftung (2020) finden sich ein ausführlicher Überblick zum Produktivitätsbegriff und seinen Entwicklungsfaktoren sowie ein umfangreiches Literaturverzeichnis mit den wichtigsten Quellen.

Herkömmliche Erklärungen des verlangsamten Wachstums der Totalen Faktorproduktivität

Kompositionseffekt

Das verlangsamte TFP-Wachstum ist zunächst eine zwangsläufige Begleiterscheinung, die in Deutschland untrennbar mit dem „German Job Miracle“[16] einhergeht – also der immensen Beschäftigungserfolge der letzten fünfzehn Jahre. Zwischen 2005 (mit einer jahresdurchschnittlichen Erwerbspersonenzahl von rund 39 Millionen) und 2020 (mit einer Erwerbspersonenzahl von rund 45 Millionen vor Ausbruch der Coronapandemie im ersten Quartal 2020) sind in Deutschland rund sechs Millionen zusätzliche Erwerbstätige in den Arbeitsmarkt integriert worden.[17]

Kehrseite des Anstiegs der Erwerbsverhältnisse ist die plausible Annahme, dass die neu beschäftigten Personen zuvor (als Massen von auch vergleichsweise gut Qualifizierten Arbeit suchten) ihre Fähigkeiten nicht erfolgreich auf dem Arbeitsmarkt einbringen konnten, „sodass davon auszugehen ist, dass sie insgesamt weniger produktiv waren als der durchschnittliche Erwerbstätige im Jahr 2005.“[18] Mit dem Zugang von weniger produktiven Erwerbspersonen in den Arbeitsmarkt dürfte die durchschnittliche Produktivität gesunken sein, was als „Kompositionseffekt“ bezeichnet wird.[19]

Der „Kompositionseffekt“ erklärt und relativiert auch etwas die Folgen der Arbeitsmarktreformen seit Mitte der Nullerjahre. Die Politik des „Förderns und Forderns“ war darauf ausgerichtet, möglichst viele Personen in Arbeit zu bringen. Arbeit sollte sich mehr lohnen, als arbeitslos zu bleiben. Entsprechend wurde das Anrecht auf staatliche Unterstützung verkürzt und verschärft, und wer arbeitslos war, musste auch vergleichsweise schlechter bezahlte Jobs akzeptieren. Insbesondere wurde die Beweislast umgekehrt. Nicht mehr die Arbeitsagentur

---

[16] Economist (2013).

[17] Daten aus IAB (2020).

[18] Sachverständigenrat (2015), S. 288.

[19] Sachverständigenrat (2016), S. 129. Der „Kompositionseffekt“ ist – mit gegenteiligem Vorzeichen – das Gegenstück zur „Entlassungsproduktivität“, die entsteht, wenn eine gegebene Wertschöpfung mit deutlich weniger Erwerbstätigen erbracht wird, was die (Arbeits-)Produktivität ansteigen lässt. Diese „Verwandtschaft“ erklärt auch sich auf den ersten Blick scheinbar widersprechende empirische Ergebnisse – siehe dazu die nachfolgenden Beiträge von Roth (2021) und von Schneider (2021) in diesem Sammelband.

musste nachweisen, dass bestimmte Tätigkeiten zumutbar waren. Vielmehr lag es nun an den Arbeitslosen, zu belegen, wieso ihnen die Annahme eines offenen Stellenangebots unzumutbar sein sollte. Im Ergebnis nahm die Beschäftigung in Deutschland rasant zu, aber bei vergleichsweise nur moderat steigenden Löhnen.

Eine lange Periode der Lohnzurückhaltung – wie sie für eine gewisse Zeit in Deutschland typisch war – muss jedoch nahezu zwangsläufig mit niedrigen Produktivitätszuwächsen einhergehen.[20] Die Logik hinter dieser Erwartung ist simpel: Wenn Arbeit billig(er) ist, unterbleiben betriebswirtschaftliche Anreize, in Maschinen, Roboter und neue digitale Technologien zu investieren. Warum auf teure Automaten setzen, wenn Arbeit so billig ist? Entsprechend unterbleibt ein Modernisierungsschub. Es wird mit billiger Arbeitskraft zu arbeitsintensiv, mit zu viel Arbeit und zu wenig Kapital produziert, was eben dazu führt, dass die Arbeitsproduktivität eben nur vergleichsweise langsam ansteigt. Anstelle von Robotern werden Menschen beschäftigt, anstatt künstliche Intelligenz einzusetzen, werden Personen beansprucht. Das Ergebnis: viele Jobs, aber wenig Innovation und somit geringe Arbeitsproduktivität und schlechte Bezahlung.[21]

Strukturwandel

Eine Verlangsamung der Produktivitätsentwicklung kann ebenso das Ergebnis einer strukturellen Verschiebung weg von hochproduktiven Wirtschaftsbereichen wie dem Verarbeitenden Gewerbe hin zu Bereichen mit geringerer Produktivität sein, wie sie in vielen Dienstleistungsbereichen zu verzeichnen ist. Das „Forellenquintett" von Franz Schubert liefert hierfür eine Veranschaulichung. Die Arbeitsproduktivität der fünf Musizierenden ist – allem technischen Fortschritt zum Trotz – auf Punkt und Komma genau dieselbe wie vor 200 Jahren bei der Uraufführung. Wie zu Schuberts Zeiten lassen sich weder Klavier, Violine, Viola oder Violoncello noch Kontrabass einsparen – alle fünf Musizierenden sind unverzichtbar – gestern genauso wie heute und morgen, neue Technologien hin oder her. Und ebenso wenig kann einer der fünf Sätze fallengelassen werden. Bestenfalls lassen sich die Pausen zwischen den Sätzen kürzen und kann etwas rascher gespielt werden. Aber insgesamt dürfte die

---

[20] Empirische Belege für diese These finden sich bei Vergeer und Kleinknecht (2014).

[21] Vgl. hierzu Kleinknecht (2017).

Spieldauer von 38 Minuten höchstens um ein paar wenige Sekunden verkürzt werden. Ein Produktivitätsfortschritt jedoch sieht völlig anders aus.

Was für die Musizierenden gilt, trifft natürlich auf eine Vielzahl von Aktivitäten im Dienstleistungswesen gleichermaßen zu. Lehrkräfte können ihre Arbeitsproduktivität prinzipiell nicht steigern. Wie auch? Wohl könnte die Anzahl der zu Unterrichtenden pro Ausbildungsstunde erhöht werden. Das würde zwar in den Statistiken als Produktivitätsanstieg verbucht. In der Realität allerdings dürfte die Qualität des Unterrichts schlechter und nicht besser werden, wenn anstelle von 20 Kindern 30 in derselben Schulklasse oder Kindergruppe sitzen.

Gleiches zeigt sich im Gesundheitswesen. Da werden Algorithmen die Produktivität der Ärztinnen und Ärzte bei Problemerkennung, Diagnose und Therapie gewaltig steigern, weil viel Bürokratie, Wiederholung, Papierkram entfallen und künstliche Intelligenz menschliches Suchen nach Ursachen und Wirkungen verbessert. Aber bereits in der Betreuung und Beratung, in der Krankenpflege und erst recht in der Altenpflege oder Psychiatrie wird es schwieriger, persönliche Anteilnahme, Empathie oder Seelsorge durch Roboter und Maschinen zu ersetzen.

Während in Bürokratie und Verwaltung durch digitales Datenmanagement sogar ein enormer Produktivitätssprung möglich wäre, ist bei Polizei und Notfalldiensten eine weitere Personaleinsparung kaum mehr möglich, ohne die Qualität der Daseinsvorsorge und des öffentlichen Angebots zu verschlechtern. Das dürfte auch bei vielen staatlichen, kommunalen oder kirchlichen Stellen wie Sozialeinrichtungen, Tagesstätten und Kinderheimen oder im Justiz(vollzugs-)wesen, im Kulturbereich und bei Freizeit und Sport zutreffen. Im Fußball waren es immer 22 Spielende und die Anzahl der Offiziellen ist gestiegen und nicht etwa gesunken, gerade auch neuer Technologien wie beispielsweise des Videobeweises wegen.

Das Beispiel des Forellenquintetts veranschaulicht, dass mit dem Strukturwandel von der kapitalintensiven Industrie- zur arbeitsintensiven Dienstleistungsgesellschaft die gesamtwirtschaftliche (Arbeits-)Produktivität zwangsläufig langsamer vorankommt. Die im höchsten Fall geringe Steigerungsfähigkeit der Arbeitsproduktivität gehört zum Wesen vieler Dienstleistungen. Und noch etwas wird ersichtlich, wenn man sich klar macht, dass die Produktivitätsmessung zwei Komponenten verknüpft.

Zum einen geht es bei der Arbeitsproduktivität (und deren Entwicklung) um das Reale, also darum, was tatsächlich geschaffen, produziert oder musiziert wird. Da hat sich beim Forellenquintett in 200 Jahren gar nichts getan. Die Statistik kann hier nur ewige Stagnation messen.

Dann jedoch geht es um das Monetäre, also darum, was Produkte, Leistungen und Musikgenuss kosten. Musikliebhaber sind heute bereit, ein Vielfaches dessen zu zahlen, was zu Beginn des 19. Jahrhunderts üblich war, obwohl das Forellenquintett um kein Jota anders ist als damals. Ganz offenbar wird heutzutage der Wert eines Konzertbesuchs höher eingeschätzt als vor 200 Jahren. Gründe für die höhere Zahlungsbereitschaft der Kunden finden sich in geänderten Vorlieben, daneben aber vor allem auch im allgemein höheren Wohlstand – was nun wiederum eine Folge dessen ist, dass sich in der Industrie und bei spezialisierten Dienstleistungen durchaus der technische Fortschritt positiv bemerkbar macht, was das Lohnniveau allgemein ansteigen lässt. Im Ergebnis kommt es in der Statistik zu einer indirekten und damit nur bescheidenen Zunahme der Arbeitsproduktivität auch der Musikerinnen und Musiker, die das Forellenquintett darbieten – selbst wenn das Reale völlig unverändert dasselbe geblieben ist.

Spätestens hier wird deutlich, dass die Verhältnisse in Deutschland nicht durch ein Produktivitätsparadox, sondern durch ein Dienstleistungsparadox gekennzeichnet sind. Solange sich hierzulande die Wertschätzung für eine gute Qualität der Dienstleistungen nicht in entsprechender Bezahlung ausdrückt, solange also im Gesundheits-, Bildungs- und Sozialwesen, bei öffentlichen Dienstleistungen oder in Kunst und Kultur die Löhne nicht stärker steigen, wird die Arbeitsproduktivität im Kriechgang oder überhaupt nicht vorankommen. Und wenn sich da nichts ändert, werden Roboter und künstliche Intelligenz das scheinbare Produktivitätsparadox vergrößern und nicht verkleinern, da sie die Arbeitsproduktivität für wenige erhöhen, nicht jedoch für die Masse der Dienstleistenden. Wirkliche Abhilfe bringt da deshalb nur eines: Dienstleistungen so hoch zu schätzen, wie es für eine auf Qualität setzende Gesellschaft unabdingbar ist.

Digitalisierung und immaterielle Wertschöpfung

Mit der strukturellen Transformation von der physischen Industrie- zur güterlosen Dienstleistungs- und nun zur digitalen Wissens- und Datenökonomie ändert sich das Wesen der Ökonomie fundamental. Insbesondere führte die digitale Transformation der vergangenen Dekaden zu einer steigenden Bedeutung immaterieller Produktionsfaktoren.[22] Der Anteil von Investitionen in immateri-

[22] Vgl. hierzu Corrado et al. (2009) und (2012) sowie den Beitrag von Roth (2021) in diesem Band.

elle Produktionsfaktoren an der Bruttowertschöpfung stieg zwischen den Jahren 1995 und 2016 in Deutschland und den USA um etwa 2 Prozentpunkte und in Frankreich um etwa 4 Prozentpunkte an.[23]

Zu den immateriellen Sachverhalten, die mit Digitalisierung und Datenökonomie einhergehen, gehören auch mikroökonomische Prozesse, die mit der Adoption von und Adaption an neue Technologien zu tun haben. Dazu gehören insbesondere auch auf andere Bereiche überschwappende Wirkungen (spillovers), die als positive externe Effekte Prozesse und Abläufe verbessern.

Das Hauptproblem der Zunahme der immateriellen Wertschöpfung liegt darin, dass deren Beitrag zur gesamtwirtschaftlichen Wertschöpfung nicht gemessen, sondern nur geschätzt werden kann. Damit jedoch verschiebt sich deren statistische Erfassung von der Realität in die Welt abstrakter Modelle. Es sei noch einmal daran erinnert, dass das Wachstum der TFP in aller Regel als Restgröße (als sogenanntes Solow-Residuum) berechnet wird, und zwar als nicht direkt erklärbare Wertschöpfung, die sich ergibt, „nachdem vom berechneten Produktionswachstum die gemessenen Veränderungen der einzelnen Produktionsfaktoren, gewichtet mit ihren Produktionselastizitäten, abgezogen wurden.“[24] Wie anfällig derartige residuale Schätzverfahren für Fehler aller Art sind, dürfte offensichtlich sein.[25]

Daten sind anders zu messen als Waren

Daten sind wie Dienstleistungen und Wissen nicht dinglicher Natur. Das macht es für die Statistik so schwierig, sie zu erfassen. Deshalb tun sich herkömmliche Verfahren zur Messung ökonomischer Aktivitäten so schwer. Daten wabern virtuell und „undinglich“, losgelöst von Standorten im In- und Ausland durch Internet und Orbit. Eine geografische Abgrenzung von Herkunft, Verarbeitung, Vermarktung und Nutzung wird dadurch nahezu vollständig unmöglich gemacht. Information und Kommunikation beruflicher Art, aber auch innerhalb sozialer Netzwerke, lassen sich nur schwerlich an nationalen Grenzen aufhalten, auch nicht, wenn Zollstationen errichtet oder Schutzmauern verstärkt werden.

---

[23] Sachverständigenrat (2019), S. 165.

[24] Sachverständigenrat (2019), S. 90.

[25] Vgl. dazu Jacobs und van Norden (2015), zitiert nach Sachverständigenrat (2015), S. 284.

Daten sind um Dimensionen rascher hin und her zu bewegen als Güter, aber auch als Dienstleistungen, bei denen sich entweder Patienten in die Praxis oder Ärzte zu den Patienten nach Hause begeben müssen. Bei Onlinebehandlungen erfolgt eine Beratung aus der Ferne. Telemedizin erlaubt, Operationen durchzuführen, bei denen die Chirurgin Tausende von Kilometern weit entfernt vom Patienten das Skalpell fernsteuert. Künftig werden mit Datenbanken verbundene implantierte Sensoren und mithilfe von Algorithmen dafür sorgen, dass Herzschrittmacher, Insulinpumpen, Blutdruckregler und vieles andere mehr automatisch und ohne das Zutun Dritter das auslösen, was medizinisch geboten scheint. Dass sich damit eine Reihe von juristischen Verantwortlichkeits- und Haftungsfragen völlig neu stellt, ist offensichtlich. Genauso klar wird, dass digital erbrachte Dienstleistungen eine neue Form der statistischen Erfassung erforderlich machen und insbesondere (räumliche) Abgrenzungsprobleme zu bewältigen sind.

Messfehler als Erklärung

Mit den rasanten Veränderungen von Datenwirtschaft und Digitalisierung vermag das BIP nicht mitzuhalten. Das Bruttoinlandsprodukt (BIP) als heutzutage in Ökonomie und Ökonomik unverändert mit Abstand wichtigster Maßstab zur Messung wirtschaftlicher Aktivitäten ist auf Volkswirtschaften ausgerichtet, die sich nach Maßgabe von Landesgrenzen trennen lassen. Erfasst werden Raum und Zeit zurechenbare, territorial nach Gemeinden, Regionen oder Staaten abgrenzbare wirtschaftliche Tätigkeiten, die Werte schaffen oder für Einkommen oder Umsätze sorgen. Bereits beim Einbezug „öffentlicher" Güter (für die es keine Marktpreise gibt), muss mangels direkter Messverfahren auf indirekte Hilfsberechnungen ausgewichen werden (was gerade wegen des steigenden „öffentlichen" Charakters der Wertschöpfung innerhalb von Digitalisierung und Datenökonomie ein zunehmendes Problem darstellen dürfte).

Die Wertschöpfung der Datenökonomie wird nur rudimentär einbezogen.[26] Deshalb sind die Mängel des BIP und der Verfahren zu seiner Messung im Zeitalter von Digitalisierung, Datenwirtschaft, Algorithmen und künstlicher

[26] Brynjolfsson und Saunders (2009) sowie Brynjolfsson und Oh (2012). Brynjolfsson et al. (2019) entwickeln mit dem GDP-B einen Ansatz, mit dessen Hilfe sich der Nutzen (und eben nicht Umsätze oder Kosten wie beim herkömmlichen BIP-Konzept) der Digitalwirtschaft und Datenökonomie mikroökonomisch messen und makroökonomisch hochrechnen lässt.

Intelligenz bedeutender denn je.[27] Weder Entdinglichung, Entterritorialisierung noch Entstaatlichung findet sich in den gängigen Statistiken zur Erfassung wirtschaftlicher Aktivitäten wieder. Das BIP ist in keiner Weise geeignet, den zunehmenden Stellenwert der Datenwirtschaft oder hybrider digitaler Staatlichkeit oder immaterieller Wertschöpfung in adäquater Weise abzubilden. Deshalb dürften die BIP-Statistiken die tatsächliche Entwicklung unterschätzen.[28]

Wie soll das BIP, das Maß aller statistischen Dinge, das Unmessbare messbar machen? Wenn die Zukunft so anders als die Vergangenheit sein wird, Dinge sich so schnell und so fundamental ändern, wie das mit den durch Digitalisierung und Datenökonomie verursachten Veränderungen der Fall sein dürfte, wenn viele Entwicklungen heutzutage noch unsicher sind und unsichtbar erfolgen, weil sie sich in den virtuellen Clouds des weltweiten Internet abspielen oder – und viel banaler – wenn sich Austausch und Handel der statistischen Erfassung entziehen, weil vieles nicht (mehr) über Märkte gekauft, sondern zwischen Haushalten geteilt und getauscht wird, wie es typisch ist für eine Sharing Economy, in der es um die Nutzung und nicht das Besitzen geht – wenn das alles mehr oder weniger gleichzeitig passiert, dann entstehen mit Statistiken und Maßzahlen, die für eine ganz andere Welt vergangener Zeiten entwickelt wurden, echte Messprobleme.

Für viele mit der von Raum und Material losgelösten Neuerungen der Digitalisierung fehlen schlicht (noch) die gesamtwirtschaftlichen statistischen Maßzahlen. Informationsgüter mit Netzwerkcharakter werden bestenfalls teilweise erfasst. Wenn moderne Apps oder Plattformen wie Uber, Car to Go oder Airbnb eine Sharing Economy und damit eine weitaus effizientere Ausnutzung

---

[27] Die statistischen Ämter behelfen sich bei der BIP-Ermittlung, in dem sie den Wert der durch die Digitalisierung geschaffenen Wertschöpfung oder den durch digitale Güter und Dienstleistungen entstehenden Nutzen oder Umsatz mit Plausibilitätsüberlegungen schätzen und nach neuen Verfahren suchen (vgl. dazu Ahmad/Ribarsky/Reinsdorf [2017] sowie Ahmad/Schreyer [2016]). So aber verlagert sich beim BIP die Erfassung noch weiter weg von der Marktorientierung hin zur Modellwelt der Statistik. Damit aber steigt die Gefahr, mit künstlich geschaffenen Zahlen den Bezug zur „gemessenen“ Realität zu verlieren. Das gilt natürlich – bei allen Vorteilen – auch für das GDP-B-Konzept von Brynjolfsson et al. (2019).

[28] Richtig ist jedoch auch, dass „ein Teil der neuen Güter nicht in den Volkswirtschaftlichen Gesamtrechnungen erscheint, die negative Substitutionseffekte jedoch dort voll sichtbar sind” (Grömling 2016), S. 139. So steigt die im BIP nicht abgebildete Konsumentenrente an, da Menschen digitale Dienste zu Preisen weit unter ihrer Zahlungsbereitschaft erhalten. Ebenso wird ein Teil dieser realen Steigerung der Kaufkraft der Haushalte durch eine nahezu kostenlose Übertragung eigener persönlicher Daten finanziert, was wiederum bei den Firmen nicht als (Datenbeschaffungs-)Kosten verbucht wird.

vorhandener Güter, Autos oder Wohnungen ermöglichen, bildet das BIP die Auswirkungen, wenn überhaupt, so nur mit rudimentären Näherungen ab.[29]

Eine beachtliche Fülle von Konsumgütern wird im Internet nahezu kostenlos zur Verfügung gestellt. Alles, was keine Umsätze erzielt, erscheint jedoch nicht im BIP. Vieles davon ersetzt aber den Kauf alternativer Marktprodukte, die im BIP erfasst waren. Wenn Wikipedia den Zugriff zu einem Onlinelexikon und YouTube das Abspielen von Videos und Filmen ermöglicht oder wenn Nerds ihre selbstentwickelten Spiele, Softwares, Musikvideos oder Fotos für die Allgemeinheit im Netz anbieten, dann erhalten die Nutzerinnen und Nutzer etwas, ohne dafür nennenswert bezahlen zu müssen. Aber obwohl die Zufriedenheit oder das Wohlbefinden der Menschen zunimmt, fällt das BIP, weil weniger Lexika oder DVDs gekauft werden. Gleiches gilt, wenn kostenpflichtige Printmedien durch frei zugängliche elektronische Nachrichtenportale ersetzt werden. Dann erhalten Menschen billiger, schneller und einfacher Zugang zu Informationen. Das BIP jedoch sinkt, weil traditionelle Medien wie Zeitungen und Zeitschriften Umsatzeinbußen erleiden, was zu Entlassungen und geringerer Wertschöpfung führt.[30]

Schlussfolgerungen: Neuvermessung der Datenökonomie

Die Wertschöpfung im Internet, der virtuelle Handel mit digitalen Daten sowie die Effekte einer Sharing Economy, in der gerade teure und langlebige Güter – wie Wohnungen, Autos oder Elektrogeräte – gemeinsam genutzt und nicht selbst erworben werden, entziehen sich in beachtlichen Teilen der sachlichen Erfassung, räumlichen Zuordnung und zeitlichen Abgrenzung. Deshalb sind das BIP und seine Messverfahren von geringerer Aussagekraft denn je. Das zu erkennen, ist nicht nur eine minimale Anforderung an die Wissenschaft, die nach neuen und besseren Methoden zu suchen hat. Es muss auch Öffentlichkeit und Medien vermittelt werden, wie unsinnig es geworden ist, sich beim BIP um marginale Schwankungen hinter dem Komma Gedanken zu machen oder gar Sorgen oder Euphorie zu verbreiten, wenn grundsätzliche Probleme das ganze Konzept in Frage stellen.

---

[29] Vgl. hierzu Demary (2015) sowie Zervas/Proserpio/Byers (2017).

[30] Zu dem Produktivitätsparadoxon, dass sich das BIP zunächst abschwächt, bevor es später zunimmt, vgl. Triplett (1999) sowie Grömling (2016) und Van Ark (2016).

„Measuring the Digital Transformation“ lautet der Titel einer aktuellen OECD-Publikation, die danach strebt, aufzuzeigen, wie sich der Wert von Daten messen und in die gängigen Statistiken zur Erfassung ökonomischer Aktivitäten einbinden lässt. Quintessenz der OECD-Analyse ist, dass, „as existing metrics and measurement tools struggle to keep up, it is imperative to act now. … More must be done to strengthen the evidence base needed to monitor and shape the digital transformation.”[31]

In welche Richtung sich neue Messkonzepte entwickeln sollten, um einer digitalisierten Datenökonomie Rechnung zu tragen, zeigen die neuen Datenquellen und Methoden, die das Statistische Bundesamt seit Anfang der 2020er-Jahre unter der Rubrik „Experimentelle Daten“ zusammenfasst.[32] So werden nun Umsatzsteuervoranmeldungen, Kreditvergaben und Auskünfte für Onlinetransaktionen oder Mobilfunkdaten (zur Ermittlung eines tagesaktuellen Mobilitätsindikators) erfasst. Aus Satellitendaten, die das Schiffsaufkommen und die Zahl der Container in Häfen aufdecken, lassen sich Rückschlüsse in Echtzeit auf Handelsaktivitäten und Produktionszahlen ziehen. Genauso können Satellitenbilder zeigen, wie dicht Parkplätze vor Geschäften belegt sind, was tagesaktuelle Erkenntnisse zu Konsumverhalten und Umsatzzahlen ermöglicht. Das Bundesamt für Güterverkehr hat aus digitalen Prozessdaten der Lkw-Mauterhebung einen Lkw-Maut-Fahrleistungsindex entwickelt, der frühzeitig (innerhalb einer guten Woche) Anhaltspunkte zur Entwicklung der Industrieproduktion liefert.

Desgleichen veröffentlicht die Deutsche Bundesbank mit dem wöchentlichen Aktivitätsindex (WAI) einen Index, der die realwirtschaftliche Aktivität in Deutschland möglichst zeitnah messen soll.[33] Er basiert auf Stromverbrauch, Lkw-Fahrleistungen, Flugbewegungen, Google-Suchanfragen, Passantenaufkommen, Luftverschmutzung und Verbraucherbefragungen.

Mit diesen modernen und Big Data angemessenen Experimenten sollen Echtzeitentwicklungen rascher sichtbar gemacht werden. Aber so begrüßenswert die Initiativen des Statistischen Bundesamts und der Deutschen Bundesbank auch sind, so sehr kratzen sie bestenfalls an der Oberfläche eines statistischen Eisbergs, dessen massiger Kegel weiterhin in der Dunkelheit der Nichterfassung liegt. Onlinemarktplätze für Unterkunft (Vermietung, Haus-

[31] OECD (2019), S. 16.

[32] Statistisches Bundesamt (2020).

[33] Deutsche Bundesbank (2020).

tausch), Transport (Mitfahrgelegenheit, Carsharing, Parkplatzsharing), Onlinearbeitsmärkte für ungelernte und qualifizierte Arbeit (Zustellung und Haushaltsdienstleistungen, IT-Berufe, Beratung) sowie (Crowdfunding) und Kreditplattformen erfordern völlig andere Messverfahren als im Fall von Güterproduktion und Warenhandel, um ökonomische Aktivitäten und Transaktionen statistisch sichtbar zu machen.

Ohne eine fundamentale Anpassung von Statistiken droht der Wirtschaftspolitik zunehmend ein Stochern im Nebel der Unschärfe und Ungenauigkeit. Das wird dann dramatisch, wenn die positiven Folgen neuer Technologien wie beispielsweise der Digitalisierung – da nicht klar erkennbar – unterschätzt und somit Produktivitätsfortschritte viel zu gering ausgewiesen werden, was wiederum die Reallohnentwicklung negativ beeinflusst. In Folge dessen wird das reale Wirtschaftswachstum unterschätzt, was sich negativ auf Stimmungen und Erwartungen und damit auf Konjunkturprognosen auswirkt.

Wenn unvollständig, inkorrekt und irreführend gemessen wird, werden auch Risiken falsch eingeschätzt, Effekte unternehmerischer oder wirtschaftspolitischer Entscheidungen über- oder unterschätzt und bei der Bewertung von Alternativen und deren Vor- und Nachteilen unzutreffende Urteile gefällt. Um Messfehler künftig zu verringern und im besten Fall zu verhindern und das Wesen einer Wirtschaft der Entdinglichung, Entterritorialisierung und Entstaatlichung abzubilden, sollten Wirtschaft, Politik und Gesellschaft mehr denn je Abstand nehmen vom alten BIP-Konzept. Stattdessen bedarf es eines Neuverständnisses in Ökonomie und Ökonomik und darauf aufbauend einer Neuvermessung wirtschaftlicher Aktivitäten und Transaktionen.

Literatur

Ahmad, Nadim; Ribarsky Jenifer; Reinsdorf, Marshall (2017): Can Potential Mismeasurement of the Digital Economy Explain the Post-Crisis Slowdown in GDP and Productivity Growth? In: OECD Statistical Working Paper No.85. OECD, Paris.

Ahmad, Nadim und Schreyer, Paul (2016): Measuring GDP in a Digitalised Economy, OECD Statistical Working Paper No. 85. OECD, Paris.

Askenazy, Philippe; Bellmann, Lutz; Bryson, Alex; Galbis, Eva Moreno (Hrsg.; 2016): Productivity Puzzles Across Europe. Oxford University Press, Oxford.

Bertelsmann Stiftung (Hrsg.; 2020): Produktivität und inklusives Wachstum. Verlag Bertelsmann Stiftung, Gütersloh.

Brynjolfsson, Erik; Collis, Avinash; Diewert, W. Erwin; Eggers, Felix; Fox, Jevin J. (2019): GDP-B: Accounting for the Value of New and Free Goods in the Digital Economy. Cambridge, Mass.: NBER Working Paper 25695. Online abrufbar unter https://www.nber.org/papers/w25695, zuletzt aufgerufen am 13.4.2021.

Brynjolfsson, Erik und Oh, Joo Hee (2012): The Attention Economy. Measuring the Value of Free Digital Services on the Internet. Online abrufbar unter https://core.ac.uk/reader/301358781, zuletzt aufgerufen am 1.6.2021.

Brynjolfsson, Erik und Saunders, Adam (2009): What the GDP Gets Wrong (Why Managers Should Care). In: MIT Sloan Management Review, Vol. 51, Nr. 1, S. 95–96. Online abrufbar unter https://sloanreview.mit.edu/article/what-the-gdp-gets-wrong-why-managers-should-care, zuletzt aufgerufen am 13.4.2021.

Byrne, David M.; Fernald, John G. ; Reinsdorf, Marshall B. (2016): Does the United States Have a Productivity Slowdown or a Measurement Problem? In: Brookings Papers on Economic Activity, Spring, S. 109–157. Online abrufbar unter https://www.brookings.edu/wp-content/uploads/2016/03/byrnetextspring16bpea.pdf, zuletzt aufgerufen am 13.4.2021.

Corrado, Carol; Haskel, Jonathan; Jona-Lasinio, Cecilia; Iommi, Massimiliano (2012): Intangible Capital and Growth in Advanced Economies. Measurement Methods and Comparative Results. IZA DP No.6733. IZA – Institute of Labor Economics, Bonn. Online abrufbar unter https://ftp.iza.org/dp6733.pdf, zuletzt aufgerufen am 13.4.2021.

Corrado, Carol; Hulten, Charles; Sichel, Daniel (2009): Intangible Capital and US Economic Growth. In: Review of Income and Wealth 55(3), 661–685.

Coyle, Diane (2014): GDP. A brief but affectionate history, Princeton University Press New Jersey.

Demary, Vera (2015): Mehr als das Teilen unter Freunden. Was die Sharing Economy ausmacht. In: Wirtschaftsdienst, 95. Jg., 2015, H. 2, S. 95–98.

Deutsche Bundesbank (2020): Wöchentlicher Aktivitätsindex für die deutsche Wirtschaft. Online abrufbar unter https://www.bundesbank.de/de/statistiken/konjunktur-und-preise/woechentlicher-aktivitaetsindex, zuletzt aufgerufen am 13.4.2021.

Economist (2013): Dissecting the miracle. Special Report on Germany. 13. Juni 2013. Online abrufbar unter https://www.economist.com/special-report/2013/06/13/dissecting-the-miracle, zuletzt aufgerufen am 13.4.2021.

Einav, Liran und Levin, Jonathan (2014): Economics in the age of big data. In: Science, Vol. 346, Issue 6210, 7 November, S. 715. Online abrufbar unter https://science.sciencemag.org/content/346/6210/1243089.long , zuletzt aufgerufen am 13.4.2021.

Gordon, Robert J. (2016): The Rise and Fall of American Growth. The U. S. Standard of Living Since the Civil War. Princeton University Press, New Jersey.

Gordon, Robert J. (2014): The Demise of U.S. Economic Growth. Restatement, Rebuttal, and Reflections. National Bureau of Economic Research Working Papers no. 19895. Online abrufbar unter https://www.nber.org/papers/w19895, zuletzt aufgerufen am 13.4.2021.

Grömling, Michael (2016): Digitale Revolution – eine neue Herausforderung für die Volkswirtschaftlichen Gesamtrechnungen? In: Wirtschaftsdienst, 96. Jg., H. 2, S. 135 –139.

Hansen, Alvin H. (1938): Full Recovery or Stagnation? W.W. Norton & Co,New York.

IAB (2020): Daten zur kurzfristigen Entwicklung von Wirtschaft und Arbeitsmarkt (Stand: November 2020). Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit (IAB). Online abrufbar unter http://doku.iab.de/arbeitsmarktdaten/Aktuelle_Daten.pdf, zuletzt aufgerufen am 13.4.2021.

Jones, Charles (2016): The Facts of Economic Growth. In: Taylor, John B. und Uhlig, Harald (Hrsg.): Handbook of Macroeconomics, Bd. 2A. Elsevier, Amsterdam, S. 3–69.

Kleinknecht, Alfred (2017): Angebotsökonomie: wenig Innovation – viele Jobs! In: Wirtschaftsdienst, Jg. 97, H. 13, S. 25–27. Online abrufbar unter https://www.wirtschaftsdienst.eu/inhalt/jahr/2017/heft/13/beitrag/angebotsoekonomie-wenig-innovation-viele-jobs.html, zuletzt aufgerufen am 13.4.2021.

Organisation for Economic Cooperation and Development (OECD; 2019): Measuring the Digital Transformation. A Roadmap for the Future. Paris (OECD Publishing). Online abrufbar unter http://www.oecd.org/going-digital/measuring-the-digital-transformation-9789264311992-en.htm, zuletzt aufgerufen am 13.4.2021.

Roth, Felix (2021): Das Produktivitätspuzzle – eine kritische Bewertung Beitrag. In diesem Band, S. 61-82.

Royal Swedish Academy of Sciences (1987): Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1987 to Robert M. Solow. Press Release of 21 October 1987. Online abrufbar unter https://www.nobelprize.org/prizes/economic-sciences/1987/press-release, zuletzt aufgerufen am 13.4.2021.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftichen Entwicklung (2020): Corona-Krise gemeinsam bewältigen, Resilienz und Wachstum stärken. Jahresgutachten 2020/21. Online abrufbar unter https://www.sachverstaendigenrat-wirtschaft.de/jahresgutachten-2020.html, zuletzt aufgerufen am 13.4.2021.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2019): Den Strukturwandel meistern. Jahresgutachten 2019/20. Online abrufbar unter https://www.sachverstaendigenrat-wirtschaft.de/jahresgutachten-2019.html, zuletzt aufgerufen am 13.4.2021.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2016): Zeit für Reformen. Jahresgutachten 2016/17. Online abrufbar unter https://www.sachverstaendigenrat-wirtschaft.de/publikationen/jahresgutachten/fruehere-jahresgutachten/jahresgutachten-201617.html, zuletzt aufgerufen am 13.4.2021.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2015): Zukunftsfähigkeit in den Mittelpunkt. Jahresgutachten 2015/16. Online abrufbar unter https://www.sachverstaendigenrat-wirtschaft.de/publikationen/jahresgutachten/fruehere-jahresgutachten/jahresgutachten-201516.html, zuletzt aufgerufen am 13.4.2021.

Schneider, Henrique (2021): Zwei Rätsel der Produktivität – eine empirische Beobachtung. Beitrag in diesem Band, S. 80–98.

Solow, Robert M. (1987): We'd Better Watch Out (Book Review on The Myth of the Post-Industrial Economy, by Stephen S., Cohen and John Zysman). The New York Times Book Review, 12. Juli , S. 36.

Statistisches Bundesamt (2020): Experimentelle Daten. Online abrufbar unter https://www.destatis.de/DE/Service/EXDAT/_inhalt.html, zuletzt aufgerufen am 13.4.2021.

Summers, Lawrence H. (2016): The Age of Secular Stagnation (What It Is and What to Do About It). In: Foreign Affairs, Vol. 95, Nr. 2, S. 2–9. Online abrufbar unter https://www.foreignaffairs.com/articles/united-states/2016-02-15/age-secular-stagnation, zuletzt aufgerufen am 13.4.2021.

Summers, Lawrence H. (2014): U.S. Economic Prospects: Secular Stagnation, Hysteresis and the Zero Lower Bound. In: Business Economics, Vol. 49, Nr. 2, p. 65–73. Online abrufbar unter http://larrysummers.com/wp-content/uploads/2014/06/NABE-speech-Lawrence-H.-Summers1.pdf, zuletzt aufgerufen am 13.4.2021.

Triplett, Jack E. (1999): The Solow Productivity Paradox: What do Computers do to Productivity? The Canadian Journal of Economics, Vol. 32, No. 2, Special Issue on Service Sector Productivity and the Productivity Paradox (April), S. 309–334.

Van Ark, Bart (2016): The Productivity Paradox of the New Digital Economy. In: International Productivity Monitor, No. 31, Herbst, S. 3–18.

Vergeer, Robert und Kleinknecht, Alfred (2014): Do Labour Market Reforms Reduce Labour Productivity Growth? A Panel Data Analysis of 20 OECD Countries (1960–2004). In: International Labour Review, 153. Jg. (2014), H. 3, S. 365–394. Online abrufbar unter http://alfredkleinknecht.nl/wp-content/uploads/2017/03/I-L-R-Vergeer-Kleinknecht.pdf, zuletzt aufgerufen am 13.4.2021.

Zervas, Georgios; Proserpio, Davide; Byers, John W. (2017): The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. In: Journal of Marketing Research: Vol. 54, No. 5, S. 687–705. Online abrufbar unter https://doi.org/10.1509/jmr.15.0204, zuletzt aufgerufen am 13.4.2021.

Das Produktivitätspuzzle – eine kritische Bewertung[1]

Felix Roth

Einleitung

Das Wachstum der Arbeitsproduktivität trägt maßgeblich zur Wettbewerbsfähigkeit[2] und zum steigenden Wohlstand[3] einer Volkswirtschaft bei. In den meisten fortgeschrittenen Volkswirtschaften ist es von zentraler Bedeutung, den Lebensstandard einer Gesellschaft mit alternder Bevölkerung aufrechtzuerhalten.[4] Trotz dieser immensen Bedeutung lässt sich feststellen, dass fortgeschrittene Volkswirtschaften wie die USA und die EU seit Beginn der Großen Rezession im Jahr 2007 einen deutlichen Rückgang der Wachstumsraten der Arbeitsproduktivität verzeichnet haben.[5] Nach der Finanzkrise (2008–15/16) hat sich jene im Vergleich zur Vorkrisenzeit (1995–2004/07) mehr als halbiert.[6]

Obwohl in diesen Volkswirtschaften ab den 1970er-Jahren – mit Ausnahme der USA Mitte bis Ende der 1990er-Jahre – ein stetiger Rückgang des Arbeitsproduktivitätswachstums zu beobachten ist,[7] bleibt das Ausmaß eines solchen

---

[1] Bei dem folgenden Artikel handelt es sich um die deutsche Übersetzung des englischsprachigen Artikels mit dem Titel *The Productivity Puzzle – A Critical Assessment and an Outlook on the COVID-19-Crisis* (Roth 2021). Dieser basiert auf dem Manuskript des Habilitationsvortrags *The Productivity Puzzle – A Critical Assessment*, welcher vom Autor am 29. Juni 2020 an der Universität Hamburg gehalten wurde (Roth 2020a). Der Autor möchte sich bei Thomas Straubhaar und Henning Vöpel für ihre konstruktiven Kommentare im Rahmen des gemeinsamen digitalen Workshops *Datenwirtschaft: Was ist neu und anders* am 11. November 2020 bedanken.

[2] Vgl. Krugman (1994).

[3] Vgl. Heil (2018).

[4] Vgl. Posen/Zettelmeyer (2019).

[5] Vgl. Oulton (2018), Van Ark/ O'Mahony (2016), Van Ark (2016), Van Ark / Jäger (2017), Van Ark et al. (2018).

[6] Vgl. Remes et al. (2018), Brynjolfsson et al. (2019), Van Ark et al. (2018).

[7] Vgl. Gordon (2018), Bergeaud et al. (2016), Brynjolfsson et al. (2019).

Rückgangs seit Beginn der Großen Rezession (2008–2013/16) für die im Themengebiet forschenden Wissenschaftler[8] aus zwei Gründen rätselhaft.

Zum einen sorgt jener Rückgang für Irritation innerhalb der Forschung, da die Realzinsen nahe oder unter null lagen.[9] Zum anderen ist der Rückgang rätselhaft, da er sich inmitten einer anhaltenden Revolution in der Informations- und Kommunikationstechnologie (IKT) sowie der künstlichen Intelligenz (KI) vollzog.[10] Wirtschaftswissenschaftler haben versucht, dieses Rätsel unter verschiedenen facettenreichen Bezeichnungen zu erfassen, so etwa als *The Secular Stagnation Puzzle*[11], *The Modern Productivity Paradox*[12] oder einfach als *The Productivity Puzzle*[13].

Der folgende Beitrag beginnt mit einem Überblick über ein von Roth und Thum im Jahr 2013 entwickeltes Modell für das Wachstum der Arbeitsproduktivität, welches um immaterielles Kapital erweitert worden ist.[14] Im Anschluss daran werden die Trends beim Wachstum der Arbeitsproduktivität von 1950 bis 2006 untersucht. Im dritten Teil wird auf den ausgeprägten Produktivitätsrückgang von 2007 bis 2015 eingegangen. Im vierten Teil wird das „Produktivitätspuzzle" kritisch diskutiert, indem die Schlüsselrolle immaterieller Werte, bezogen auf das Wachstum der Arbeitsproduktivität, herausgearbeitet wird. Sodann befasst sich der Beitrag mit aktuellen Fragen, die sich anlässlich der COVID-19-Krise ergeben. Zwei Hauptschlussfolgerungen bilden den Abschluss des Beitrags.

Determinanten des Arbeitsproduktivitätswachstums

Zunächst werden die Determinanten des Arbeitsproduktivitätswachstums mithilfe der Präsentation einer um immaterielles Kapital erweiterten Modellspezifikation herausgearbeitet. Letzteres wurde von Roth und Thum im Jahr 2013 im Rahmen des INNODRIVE-Projekts[15] (Intangible Capital and Innovations: Drivers

---

[8] Vgl. Oulton (2018), Remes et al. (2018), Van Ark/Jäger (2017).

[9] Vgl. Teuling/Baldwin (2014), Summers (2015), Haskel/Westlake (2018a).

[10] Vgl. OECD (2015).

[11] Summers (2014, 2015), Teulings/Baldwin (2014).

[12] Brynjolfsson et al. (2019).

[13] Haskel/Westlake (2018a).

[14] Roth and Thum (2013).

[15] INNODRIVE (2011), Roth/Thum (2013).

of Growth and Location in the EU) entwickelt und findet derzeit Verwendung in seinem laufenden GLOBALINTO-Projekt[16] (Capturing the value of intangible assets in micro data to promote the EU's growth and competitiveness).[17]

Die Modellspezifikation folgt einem Ansatz Benhabibs und Spiegels aus dem Jahr 1994,[18] der als „ländervergleichende Wachstumsbilanzierung" bezeichnet werden kann. Dieser unterscheidet sich von dem Rahmen der traditionellen Einzelwachstumsbilanzierung durch zwei Komponenten: Erstens werden die Ausgangselastizitäten geschätzt und nicht auferlegt, zweitens kann ein Teil des Modells entworfen werden, um die internationale Varianz des Wachstums der Gesamtfaktorproduktivität (GFP) zu erklären. Auf dem theoretischen Rahmen von Corrado et al. (2009)[19] aufbauend, wird die Modellspezifikation Benhabibs und Spiegels (1994)[20] um immaterielle Werte erweitert. Ausgangspunkt für die Schätzung ist sodann eine erweiterte Cobb-Douglas-Produktionsfunktion. Unter der Annahme konstanter Skalenerträge wird zunächst die Cobb-Douglas-Produktionsfunktion in intensiver Form umgeschrieben. Zweitens werden die Differenzen in den natürlichen Logarithmen genommen und der GFP-Term geschätzt. Dies bietet die folgende Grundlage für die ökonometrischen Ergebnisse, welche an späterer Stelle in diesem Beitrag veranschaulicht und diskutiert werden:

$$\left(lnq_{i,t} - lnq_{i,t-1}\right) = c + gH_{i,t} + mH_{i,t}\frac{(q_{\max,\,t} - q_{i,t})}{q_{i,t}} + n\left(1 - ur_{i,t}\right) + p\sum_{j=1}^{k} X_{j,i,t} + yd_{i,t} + \alpha\left(lnk_{i,t} - lnk_{i,t-1}\right) + \beta\left(lnr_{i,t} - lnr_{i,t-1}\right) + u_{i,t} \quad (1)$$

wo das Wachstum der Arbeitsproduktivität *($lnq_{i,t}$ – $lnq_{i,t-1}$)*, [Bruttowertschöpfung für die nicht-landwirtschaftlichen privatwirtschaftlichen Sektoren, erweitert durch die Investitionsströme von immateriellem Unternehmenskapital in Land *i* und Zeitraum *t*] im Wesentlichen in einen GFP-Term und zwei Kapitalterme – materielles und immaterielles Kapital – zerlegt werden kann. GFP wird durch einen konstanten Term *c* dargestellt, der den exogenen technologischen Fortschritt darstellt. Die Höhe des Humankapitals *($H_{i,t}$)* spiegelt die Innovationsfähigkeit eines Landes im Inland wider. Der Ausdruck $H_{i,t}\frac{(q_{max,\,t} - q_{i,t})}{q_{i,t}}$ steht für einen Aufholprozess. Der Term *($1 – ur_{i,t}$)* berücksichtigt den Konjunktureffekt.

[16] GLOBALINTO (2020).

[17] Vgl. Roth/Thum (2013), Roth (2020c).

[18] Benhabib/Spiegel (1994).

[19] Vgl. Corrado et al. (2009).

[20] Vgl. Benhabib/Spiegel (1994).

Der Ausdruck $\sum_{j=1}^{k} x_{jit}$ ist die Summe von k zusätzlichen wirtschaftspolitisch relevanten Kontrollvariablen, die möglicherweise das GFP-Wachstum erklären könnten. Diese schließen öffentliche immaterielle Vermögenswerte ein, zum Beispiel formelle und informelle Institutionen wie Rechtsstaatlichkeit und interpersonelles und systemisches Bevölkerungsvertrauen. Sie sind von zentraler Bedeutung für das Wachstum. $yd_{i,t}$ sind Jahres-Dummyvariablen, die unter anderem die wirtschaftlichen Abschwünge in den Jahren 2001 und 2008 berücksichtigen sollen. Als Nächstes folgen der Term für das Wachstum materieller Kapitaldienstleistungen $(lnk_{it}-lnk_{it-1})$ sowie der Term für das Wachstum der immateriellen Kapitaldienstleistungen $(lnr_{it}-lnr_{it-1})$ und der Fehlerterm $u_{i,t}$. Im fünften Teil dieses Beitrags werden wir auf den β-Koeffizienten für das Wachstum immaterieller Kapitaldienstleistungen eingehen.

Wachstum der Arbeitsproduktivität, 1950 bis 2006

Tabelle 1 aus von Van Ark et al. (2008)[21] weist unter anderem die durchschnittlichen jährlichen Wachstumsraten des BIP pro Arbeitsstunde in der EU-15 und den USA von 1950 bis 2006 aus. Die visualisierte empirische Evidenz zeigt, dass das Wachstum der Arbeitsproduktivität in der EU von 1950 bis 1973 mit 5,3 % doppelt so hoch war wie in den USA mit 2,5 %. Das gleiche Muster – wenn auch mit niedrigeren Zahlen – gilt für den Zeitraum 1973 bis 1995 mit Werten von 2,4 % für die EU-15 und 1,2 % für die USA. Die Forschung führt diesen Anstieg des Arbeitsproduktivitätswachstums innerhalb der EU gegenüber den USA eindeutig auf einen Aufholprozess zurück. Dieser baut auf einer starken Kompetenzbasis auf, die in der Sekundarstufe II vermittelt wird, und auf einem Produktionsprozess, der auf Nachahmung basiert. Interessanterweise ändert sich das Muster bei der Analyse des Zeitraums 1995 bis 2006. Das Wachstum der Arbeitsproduktivität innerhalb den USA stieg auf 2,3 % gegenüber 1,5 % innerhalb der EU-15.

Bei der Analyse der zugrunde liegenden Beiträge zum Wachstum der Arbeitsproduktivität in Tabelle 4 ihres Artikels stellen Van Ark et al. fest, dass dieser Rückgang des Arbeitsproduktivitätswachstums innerhalb der EU größtenteils auf eine geringer ausgeprägte wissensbasierte Wirtschaft zurückzuführen ist.[22] Eine weitere sektorale Zerlegung durch die Autoren zeigt einen deutlichen Rückgang des GFP-Wachstums in den privatwirtschaftlich organisierten Sek-

[21] Van Ark et al. (2008).

[22] Vgl. Van Ark et al. (2008).

toren der EU-15 gegenüber den USA, insbesondere bei den privatwirtschaftlich organisierten Dienstleistungen. Die Autoren führen die Produktivitätslücke bei diesen EU-Dienstleistungen auf Mängel bei die IKT ergänzenden Investitionen in immaterielles Kapital sowie auf Rigiditäten im EU-Binnenmarkt in Bezug auf Produkt-, Arbeits- und Dienstleistungsmärkte zurück.

Ähnliche Ergebnisse, die im Einklang mit diesem Gesamtargument stehen, werden von einer Gruppe von Ökonomen vorgelegt, die mit Sapir und Aghion zusammenarbeiten. Die Gruppe betont die Bedeutung öffentlicher immaterieller Vermögenswerte, nämlich die Quantität und Qualität der Hochschulbildung, um die Lücke im Wachstum der Arbeitsproduktivität zu erklären.[23] In diesem Zusammenhang kommen Brynjolfsson et al. zu dem Schluss, dass die Produktivitätslücke am besten mit dem Faktum niedrigerer Investitionen in IKT und KI und verzögerten komplementären immateriellen Kapitalinvestitionen erklärt werden könne.[24]

Das Produktivitätspuzzle, 2007 bis 2015

Dies führt zu einer Betrachtung des Zeitraums seit Beginn der Großen Rezession von 2007 bis 2015. Tabelle 1 in Van Ark et al. zeigt einen deutlichen Rückgang des Wachstums der Arbeitsproduktivität seit Beginn der Großen Rezession im Jahr 2008.[25] Die Wachstumsraten der Arbeitsproduktivität gingen im Euroraum von 1,4 % auf 0,6 % und in den USA von 2,5 % auf 1,3 %, das heißt jeweils um die Hälfte zurück. Wie Oulton (2018)[26] hervorhebt, sei ein Rückgang dieser Größenordnung außergewöhnlich und nicht nur eine Fortsetzung der historischen Trends der Vergangenheit, wie von den US-Wirtschaftshistorikern Gordon[27] und Cowen[28] behauptet. Aber was hat diesen starken Rückgang des Arbeitsproduktivitätswachstums ausgelöst?

Zwei Wirkungskanäle lassen sich als Erklärung identifizieren: Erstens wurde der Rückgang des Arbeitsproduktivitätswachstums mit einem deutlichen Rückgang des Wachstums der Gesamtfaktorproduktivität in Verbindung

---

[23] Vgl. Aghion (2008), Aghion/Howitt (2006), Aghion et al. (2007, 2008, 2010), Sapir et al. (2004).

[24] Vgl. Brynjolfsson et al. (2019).

[25] Vgl. Van Ark et al. (2018).

[26] Vgl. Oulton (2018).

[27] Vgl. Gordon (2018).

[28] Vgl. Cowen (2011).

gebracht. Die von Bergeaud et al. veranschaulichten Langzeitbefunde mit Zeitreihenbefunden zu Arbeitsproduktivitätswachstum und Gesamtfaktorproduktivität von 1890 bis 2010 stützen eine solche Behauptung.[29]

Zweitens ist der Rückgang des Arbeitsproduktivitätswachstums auf einen Rückgang der Investitionen zurückzuführen. Diese Behauptungen werden durch Analysen gestützt, welche eine erhebliche Rückläufigkeit von Investitionen in Sachkapital im Zeitraum 2008 bis 2013 aufzeigen. Der Rückgang der materiellen Investitionen innerhalb der EU-Volkswirtschaften ist in Abbildung 1 dargestellt. Jene Visualisierung orientiert sich an der Veröffentlichung von Roth aus dem Jahr 2020.[30] Insbesondere ist der stärkste Rückgang der materiellen Kapitalinvestitionen in denjenigen Peripherieländern des Euroraums beobachtbar, welche intensive Sparmaßnahmen ergriffen haben.

Dieser Rückgang des Wachstums sowie der Investitionen in die Arbeitsproduktivität hat viele Wissenschaftler vor ein Rätsel gestellt.[31] Erstens ist der Rückgang deshalb rätselhaft, weil die Realzinsen in den USA, Japan, UK und im Euroraum nahe oder unter null lagen.[32] Zweitens ist er deshalb rätselhaft, weil er inmitten einer anhaltenden Revolution in den Bereichen IKT und KI stattfand.[33] Wie Nakamura betont, deutet die Intensität der technologischen Innovationen seit Anfang der 1990-er Jahre auf eine „dramatisch dynamische Wirtschaft!" hin.[34] Wie aus Abbildung 5.4 in Haskel und Westlake hervorgeht, haben die Top-100- und Top-5 %-Frontier-Unternehmen im Gegensatz zu den Nicht-Frontier-Unternehmen seit dem Jahr 2000 tatsächlich einen enormen Anstieg des Arbeitsproduktivitätswachstums verzeichnet.[35] Darüber hinaus weist die vorhandene empirische Evidenz auf die zunehmende Bedeutung immaterieller Vermögenswerte bei den S&P-500-Unternehmen hin und ermöglicht die Feststellung, dass die zehn größten US-Unternehmen fast ausschließlich auf immateriellen Vermögenswerten basieren.[36] Eine solch aussagekräftige empirische Evidenz stützt Nakamuras Behauptung aus dem Jahre 2019[37] nachhaltig.

---

[29] Vgl. Bergeaud et al. (2016).

[30] Vgl. Roth (2020b).

[31] Vgl. Oulton (2018), Remes et al. (2018), Van Ark/Jäger (2017).

[32] Vgl. Teuling/Baldwin (2014), Summers (2015), Haskel/Westlake (2018a).

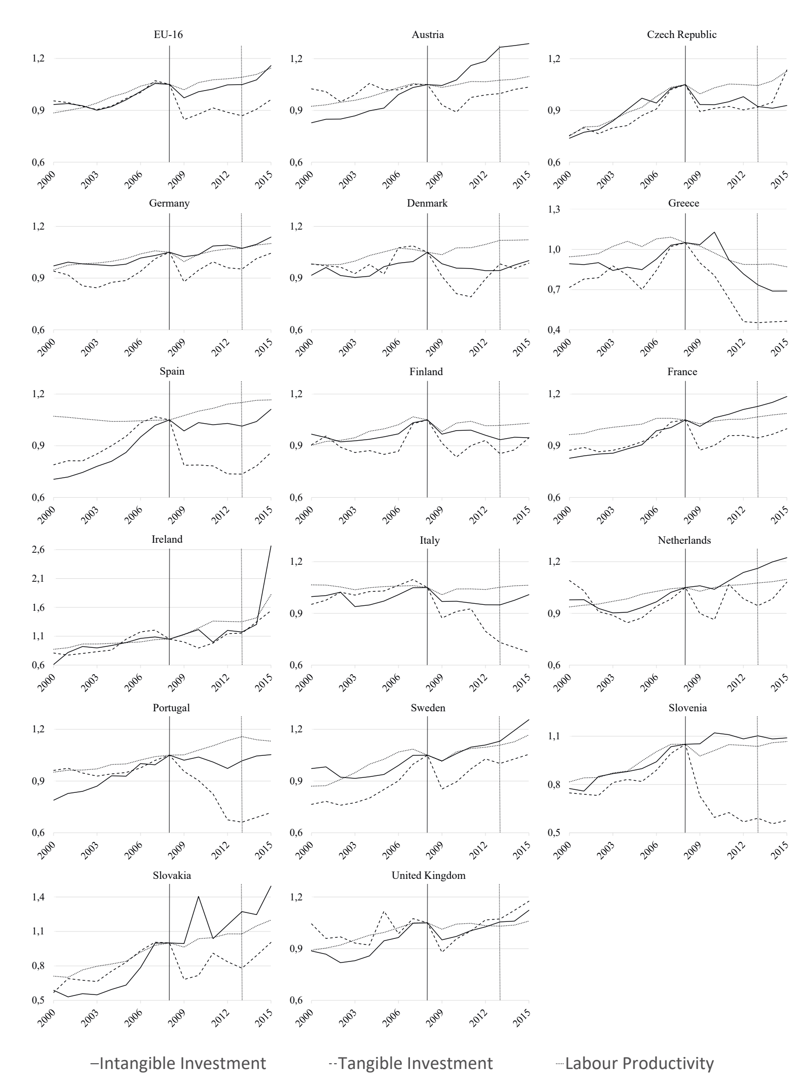
[33] Vgl. OECD (2015).

[34] Vgl. Nakamura (2019).

[35] Vgl. Haskel/Westlake (2018a), S. 95).

[36] Vgl. Ross (2020).

[37] Vgl. Nakamura (2019).



Anmerkung: Die Investitionen in immaterielles und materielles Kapital sowie die Arbeitsproduktivität sind in Millionen Landeswährungen angegeben und im Jahr 2008 auf 1 normiert. Die durchgezogene Linie zeigt den Beginn der Finanzkrise im September 2008 an. Die gestrichelte Linie zeigt den Beginn der wirtschaftlichen Erholung Ende 2013 an. Für Griechenland (Greece), Irland (Ireland) , Slovenien (Slovenia) und die Slowakei (Slovakia) werden angepasste y-Skalenwerte verwendet. Der EU-16 Durchschnitt basiert auf KKP-bereinigten Werten.

Abb. 1: Investitionen in immaterielles und materielles Kapital und Arbeitsproduktivitätswachstum, EU-16, 2000 bis 2015
Quelle: Abbildung 4 in Roth (2020b), S. 680.
Quelle(n): INTAN-Invest (NACE2) Daten Corrado et al. (2018).

Wissenschaftler wie Summers identifizieren einen Mangel an Gesamtnachfrage als Hauptursache des Rückgangs des Wachstums und der Investitionen in die Arbeitsproduktivität.[38] Die Anwendung ihrer Empfehlungen zur Stimulierung der Gesamtnachfrage innerhalb des Euroraums impliziert zwei Strategiestränge: Erstens verpflichtete sich die Europäische Zentralbank dazu, ein Programm zur quantitativen Lockerung (QE) durchzuführen. Zweitens hat sich die Europäische Kommission darauf festgelegt, einen EU-weiten Europäischen Investitionsplan auf den Weg zu bringen.[39] Dies knüpfte sie an die Bedingung, dass auf der Ebene der Mitgliedstaaten eine Strukturreformagenda verabschiedet wird, um die Grundlage für wachstumsfördernde Maßnahmen zu schaffen. Ein von einigen prominenten Ökonomen wie De Grauwe[40] und Fratzscher[41] favorisiertes Konjunkturpaket für Kernwirtschaften wie Deutschland wurde jedoch nie auf den Weg gebracht.

Dennoch war es der Politik auf EU-Ebene bereits gelungen, die Unterstützung der Nachfrage anzuregen. Wie in Abbildung 1 dargestellt, hat sie damit seit 2014 eine wirtschaftliche Erholung eingeleitet und die Investitionen im Euroraum angeregt. Die Unterstützung der Gesamtnachfrage anzuregen ist jedoch nur der erste Schritt zu einer Lösung des Produktivitätspuzzles. Ein weiterer wesentlicher Schritt ist der Einbezug immaterieller Kapitalanlagen in die Vermögensgrenze der Volkswirtschaftlichen Gesamtrechnungen.

Immaterielles Kapital und das Produktivitätspuzzle

Aber welche Investitionen in immaterielles Kapital sollten in die Vermögensgrenze der Volkswirtschaftlichen Gesamtrechnungen einbezogen werden? Corrado, Hulten und Sichel (CHS) kategorisieren in ihrem 2005 veröffentlichten wegweisenden Buchkapitel drei Dimensionen immaterieller Vermögenswerte.[42] Diese werden in Tabelle 1 dargestellt.

Zunächst führen CHS computerisierte Informationen an, welche sie als „knowledge embedded in computer programs and computerized databases" (S. 23) definieren. Als weitere Dimension nennen sie innovationsbezogenes Wis-

---

[38] Vgl. Draghi (2014), Krugman (2014), Summers (2014, 2015).

[39] Vgl. Fichtner et al. (2014).

[40] Vgl. De Grauwe (2015).

[41] Vgl. Fratzscher (2014).

[42] Vgl. Corrado et al. (2005).

sen, das sie als „not only the scienfific knowledge embedded in patents, licenses and general know-how (not patented) but also the innovative and artistic content in commercial copyrights, licenses and designs“ (S.26) verstehen. Die dritte Komponente bilden Unternehmenskompetenzen, die sie als „the value of brand names and other knowledge embedded in firm-specific human and structural resources“(S. 28) definieren.

Tab. 1: Überblick über die Definitionen der immateriellen Vermögenswerte von Unternehmen im Rahmen von CHS (2005)

CATEGORY OF INTANGIBLE ASSETS	DEFINITION BY CHS (2005)	BUSINESS INTANGIBLE ITEM	INCLUDED IN NA
Computerized information	"knowledge embedded in computer programs and computerized databases" (p.23)	Computer software	Yes
		Computerised database	Yes
Innovative property	"not only the scientific knowledge embedded in patents, licenses and general know-how (not patented) but also the innovative and artistic content in commercial copyrights, licenses and designs" (p.26)	Science and engineering R&D	Yes
		Mineral exploration	Yes
		Copyright and license costs	Yes
		Other product development, design and research expenses	No (new intangible)
Economic competencies	"the value of brand names and other knowledge embedded in firm-specific human and structural resources" (P.28)	Brand equity	No (new intangible)
		Firm-specific human capital	No (new intangible)
		Organizational structure	No (new intangible)

Anmerkung: NA = National Account.
Quelle: Eigene Adaption von CSH (2005). Tabelle I in Roth (2019), S. 6.

Inwieweit sind diese Vermögenswerte für die Stimulierung des Arbeitsproduktivitätswachstums relevant? Zwei Beispiele sind zu nennen, welche sich aus einer Kette von Argumenten ableiten, wie sie von Brynjolfsson et al.[43] in den letzten zwei Jahrzehnten entwickelt worden sind. Brynjolfsson und sein Team stellen fest, dass ein Unternehmen für jeden in Software investierten Euro zusätzliche 10 Euro für die Entwicklung von Unternehmenskompetenzen ausgeben muss, um das volle Potenzial des Wachstums der Arbeitsproduktivität auszuschöpfen. Diese zusätzlichen Investitionen umfassen die Umschulung des Personals zur Nutzung der Software sowie die Umstrukturierung der organisatorischen Abläufe. Ähnliche Ergebnisse wurden für Investitionen in KI festgestellt.

---

[43] Vgl. Brynjolfsson et al. (2000, 2002).

Welche Auswirkungen sind nun zu erwarten, wenn diese immateriellen Vermögenswerte in die Vermögensgrenze der Volkswirtschaftlichen Gesamtrechnungen einbezogen werden? Tabelle 2 aus der Arbeit von Roth aus dem Jahr 2020[44] fasst drei Hauptergebnisse zusammen.

Tab. 2: Überblick über bestehende empirische Studien, 2009 bis 2018

AUTHORS	COUNTRY	INVESTMENT (IN GDP) IN %	CONTRIBUTION TO LPG IN %†	GROWTH ACCELERATION IN %	ARTICLE	HARMONIZED CROSS-COUNTRY DATASET	METHODOLOGY
Corrado et al. (2009)	US	"~ 13* (03)"	"27 (95-03)"	"11.2 (95-03)"	RoIW	-	GA
Fukao et al. (2009)	JAP	"11.1 (00-05)"	"27; 16 (95-00); (00-05)"	"17.3, -1.4 (95-00), (00-05)"	RoIW	-	GA
Marrano et al. (2009)	UK	"13** (04)"	"20 (95-03)"	"13.1 (95-03)"	RoIW	-	GA
Nakamura (2010)	US	"Intangible=Tangible (00-07)"	/	/	RoIW	-	GA
Edquist (2011)	SE	"10/~16*** (04)"	"41; 24 (95-00); (00-06)"	"16 , -2.3 (95-00), (00-06)"	RoIW	-	GA
Roth and Thum (2013)	EU-13	"9.9**** (98-05)"	"50 (98-05)"	"4.4 (98-05)"	RoIW	INNODRIVE	CCGA
Corrado et al. (2013)	EU-15	"6.6 (95-09)"	"24 (95-07)"	/	OREP	"INTAN-Invest (NACE1)"	GA
Corrado et al. (2018)	EU-14, NMS-4	"7.2, 6.4 (00-13)"	"30, 10; 19, 8; 43‡,17 (00-13);(00-07); (07-13)"	/	JIPD	"INTAN-Invest (NACE2)"	GA

*Anmerkungen: †LPG=Labour Productivity Growth (Arbeitsproduktivitätswachstum). *Das Maß hier ist die nicht-landwirtschaftliche-Unternehmensleistung. **Das Maß hier ist die bereinigte Bruttowertschöpfung (BWS) des Marktsektors. ***Das Maß hier ist die BWS. ****Das Maß ist die BWS (ausgenommen k70). ‡Capital share (Kapitalanteil). US = United States (Vereinigte Staaten), UK = United Kingdom (Vereinigtes Königreich), JIPD = Journal of Infrastructure, Policy and Development, GA = Growth Accounting (Wachstumsbilanzierung), CCGA = Cross Country Growth Accounting (ländervergleichende Wachstumsbilanzierung). Die Zahlen in Klammern beziehen sich auf die jeweiligen Zeiträume. Quelle: Tabelle I in Roth (2020b), S. 675.*

Erstens steigen die Gesamtinvestitionen im Verhältnis zum BIP erheblich an. Die Investitionen in immaterielle Vermögenswerte nähern sich dem Niveau an, das für solche in materielles Kapital gilt, sobald immaterielle Vermögenswerte in die Vermögensgrenze der Volkswirtschaftlichen Gesamtrechnung einbezogen werden. Zweitens leisten immaterielle Vermögenswerte einen wesentlichen Beitrag zum Wachstum der Arbeitsproduktivität. So zeigt beispielsweise die Arbeit von Roth und Thum aus dem Jahr 2013, dass das Wachstum der immateriellen Kapitaldienstleistungen 50 % der internationalen Unterschiede in Bezug auf das Wachstum der Arbeitsproduktivität in der EU erklären kann.[45] Imma-

---

[44] Vgl. Roth (2020b).

[45] Vgl. Roth and Thum (2013).

terielle Kapitaldienstleistungen stellen hierbei den dominierenden Faktor dar. Drittens beschleunigt sich das Wachstum der Arbeitsproduktivität. So berichtet etwa Edquist berichtet etwa, dass sich das Wachstum der Arbeitsproduktivität in Schweden um 16 % beschleunigte, nachdem immaterielle Vermögenswerte berücksichtigt worden waren.[46]

Welche Auswirkungen haben diese Ergebnisse auf das Produktivitätspuzzle? Als Antwort auf diese Frage können vier Punkte herausgearbeitet werden: Erstens ist der „rätselhafte" Rückgang der Investitionen größtenteils auf eine fehlerhafte Messung der tatsächlichen Investitionsraten der Unternehmen in den meisten fortgeschrittenen Volkswirtschaften zurückzuführen. Die zeitgenössischen Klassifikationen der Volkswirtschaftlichen Gesamtrechnung wurden noch nicht vollständig überarbeitet, um dem anhaltenden Übergang zur wissensbasierten Wirtschaft des 21. Jahrhunderts Rechnung zu tragen. Obwohl bereits selektive Elemente immaterieller Vermögenswerte wie Software und wissenschaftliche Forschung und Entwicklung berücksichtigt wurden, sind Investitionen in Unternehmenskompetenzen wie firmenspezifisches Human- und Organisationskapital weiterhin ausgeschlossen.

Abbildung 2 aus der Arbeit von Roth aus dem Jahr 2020 zeigt, dass die Gesamtinvestitionen in eine EU-16-Länderstichprobe nach Aufnahme der immateriellen Vermögenswerte in die Volkswirtschaftliche Gesamtrechnung fast doppelt so hoch sind und 25 % der Gesamtsumme ausmachen.[47] Darüber hinaus ist es interessant zu beobachten, dass in sieben von 16 Ländern die Unternehmensinvestitionen in immaterielles Kapital bereits höher sind als in materielles Kapital.

Abbildung 1 aus der Arbeit von Roth im Jahr 2020 zeigt, dass sich die Investitionen in immaterielles Kapital trotz eines stetigen Rückgangs der Sachanlagen, insbesondere nach der Finanzkrise, rasch erholt haben und einen stetigen Aufwärtstrend verzeichnen.[48] Diese Ergebnisse decken sich mit den neuesten Erkenntnissen aus dem im Anschluss an dem INNODRIVE-Datensatz entwickelten INTAN-Invest-Datensatzes, welcher in einem Redemanuskript von Haskel aus dem Jahr 2020[49] verwendet wird. Es werden ein stetiger Rückgang des materiellen Kapitals und ein stetiger Anstieg des immateriellen Kapitals über die Finanzkrise hinaus gezeigt.

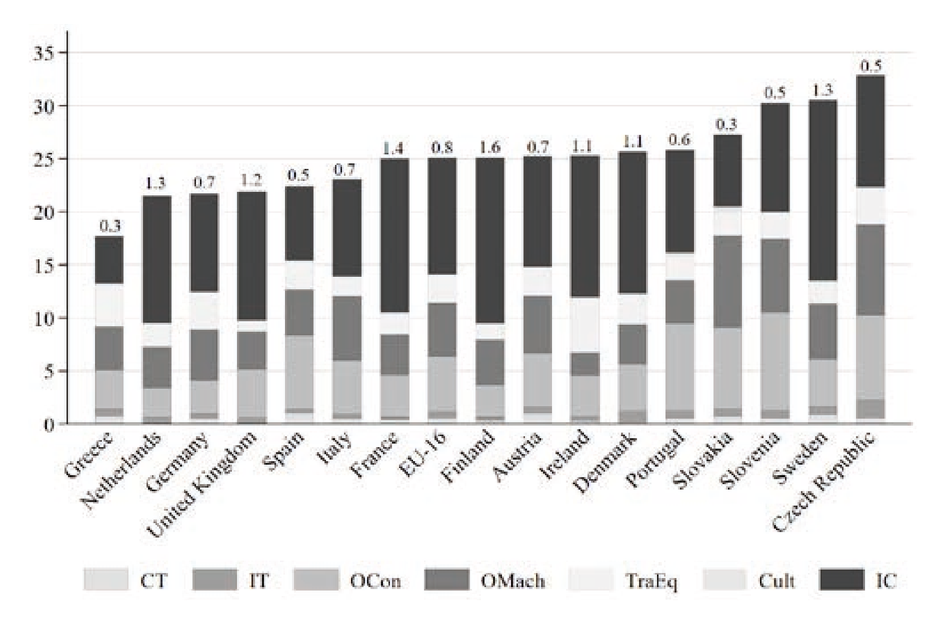
---

[46] Vgl. Edquist (2011).

[47] Vgl. Roth (2020c), S. 680.

[48] Vgl. Roth (2020b), S. 680.

[49] Vgl. Haskel (2020).



Anmerkungen: CT=communications technology (Kommunikationstechnik); IT=information technology (Informationstechnologie); OCon=total non-residential capital investment (Gesamtinvestitionen in Gewerbeimmobillien), OMach=other machinery and equipment (Andere Maschinen und Ausrüstungen); TraEQ=transport equipment (Transportausrüstung); Cult=cultivated assets (Vermögenswerte für Nutztiere und Nutzpflanzungen); IC=intangible capital (immaterielles Kapital). Wohngebäude wurden ausgeklammert. Die Werte über den Balken stellen das Verhältnis von immateriellen zu materiellen Investitionen dar.
Quelle(n): INTAN-Invest (NACE2) Daten Corrado et al. (2018) und EUKLEMS Daten Jäger (2017).

Abb. 2: Materielle und immaterielle Unternehmensinvestitionen (in Prozentpunkten der Bruttowertschöpfung), EU-16, 2000 bis 2015
Quelle: Abbildung 3 in Roth (2020b), S. 680.

Die obigen Belege zeigen, dass die Verwendung materieller Investitionsströmen als alleinige Grundlage der Analyse zu fehlerhaften empirischen Untersuchungen und letztendlich zur Implementierung fehlgeleiteter politischer Maßnahmen führt.

Zweitens bewirkt der Einbezug immaterieller Vermögenswerte in die Vermögensgrenze der Volkswirtschaftlichen Gesamtrechnungen auch eine Steigerung des Arbeitsproduktivitätswachstums. Dies hat Edquist bereits für Schweden gezeigt.[50] Seine Ergebnisse stehen hier im Gegensatz zu Behauptungen von

[50] Vgl. Edquist (2011).

Haskel und Westlake[51] und Syverson[52]. Die Ergebnisse von Roth aus dem Jahr 2020 aus dem GLOBALINTO-Projekt[53] stützen die Ergebnisse von Edquist[54]. Die Analyse von Zeiten der wirtschaftlichen Erholung zeigt, dass sich das Wachstum der Arbeitsproduktivität um 0,4 Prozentpunkte (oder 22 %) von 1,8 % auf 2,2 % beschleunigt hat. In diesem Zusammenhang prognostiziert Nakamura sogar, dass die Fehlmessung des Arbeitsproduktivitätswachstums höchstwahrscheinlich eine jährliche Wachstumsrate von 2 % ergeben werde.[55]

Drittens haben mehrere herausragende Beiträge den Zusammenhang zwischen der Höhe der Unternehmensinvestitionen in immaterielle Vermögenswerte und den Veränderungen der GFP hervorgehoben. Van Ark und O'Mahony, Van Ark und Jäger sowie Bounfour und Miyagawa führen den Rückgang der Arbeitsproduktivität und des GFP-Wachstums hauptsächlich auf eine langsamere Verbreitung von Technologie und Innovation zurück, die wiederum aus niedrigen Wachstumsraten der Investitionen in IKT und komplementäre immaterielle Werte resultiert.[56] Haskel und Westlake heben auch eine Verringerung der Spillover-Effekte immaterieller Vermögenswerte auf GFP hervor, da sich die Kluft zwischen führenden und nacheilenden Unternehmen vergrößert.[57] Darüber hinaus haben Brynjolfsson et al. argumentiert, dass mehr Investitionen in komplementäre immaterielle Vermögenswerte erforderlich sind, um den vollen Nutzen der KI für das Wachstum der Arbeitsproduktivität ausschöpfen zu können.[58]

---

[51] Vgl. Haskel/Westlake (2018a).

[52] Vgl. Syverson (2017).

[53] Vgl. Roth (2020b).

[54] Vgl. Edquist (2011).

[55] Vgl. Nakamura (2019).

[56] Vgl. Bounfour/Miyagawa (2015), Van Ark (2016), Van Ark / O'Mahony (2016), Van Ark / Jäger (2017).

[57] Vgl. Haskel/Westlake (2018b).

[58] Vgl. Brynjolfsson et al. (2019).

Tab. 3: Immaterielles Kapital und Arbeitsproduktivitätswachstum, EU-16, 2000–15, PP-PCSE Schätzung

Estimation method	PP-PCSE	PP-PCSE	PP-PCSE	PP-PCSE	PP-PCSE	2SLS
Time sample	2000-2015	2000-2015	2000-2015	2008-2015	2000-2015	2000-2015
Equation	(1)	(2)	(3)	(4)	(5)	(6)
Tangible services growth	0.31*** (0.08)	0.19** (0.08)	0.28*** (0.08)	-0.13 (0.15)	0.18** (0.07)	0.58 (0.42)
Tangible services growth*crisis	-	-	-0.32** (0.13)	-	-	-
Tangible services growth*recovery	-	-	-	0.47 (0.30)	-	-
Intangible services growth	-	0.38*** (0.07)	0.48*** (0.09)	0.32*** (0.11)	-	0.50*** (0.16)
Intangible services growth*crisis	-	-	-0.28** (0.13)	-	-	-
Intangible services growth*recovery	-	-	-	0.42* (0.23)	-	-
Innovative property services growth	-	-	-	-	0.37*** (0.07)	-
Computerized information services growth	-	-	-	-	-0.01 (0.04)	-
Economic Competencies services growth	-	-	-	-	0.02 (0.06)	-
Upper secondary education 15+	0.07*** (0.02)	0.05*** (0.01)	0.05*** (0.01)	0.02 (0.02)	0.06*** (0.01)	0.07*** (0.02)
Catch-up	-0.02** (0.01)	-0.02*** (0.01)	-0.02*** (0.01)	-0.01 (0.01)	-0.02** (0.01)	-0.02* (0.01)
Business cycle	-0.11* (0.06)	-0.12* (0.06)	-0.13** (0.06)	-0.13* (0.07)	-0.12* (0.06)	-0.11** (0.05)
R-squared	0,40	0.50	0.54	0.63	0.54	0.46
Observations	256	256	256	128	256	208
Number of countries	16	16	16	16	16	16

*Anmerkungen: In Regression (1) schließen das Wachstum der materiellen Dienstleistungen, das Wachstum der Arbeitsproduktivität und der Aufholprozess Software, F&E sowie Unterhaltung, künstlerische und literarische Originale und Mineralienexploration aus. In den Regressionen (2-6) werden das Wachstum der Arbeitsproduktivität und der Aufholprozess mit immateriellem Kapital erweitert. Materielles Kapital schließt Wohngebäude aus. Das Wachstum der Arbeitsproduktivität wurde auf der Grundlage der BWS der nicht-landwirtschaftlichen privatwirtschaftlichen Sektoren (ohne Immobilienaktivitäten) berechnet. ***p<0.01 , **p<0.05 , *p<0.1.*
Quelle: Tabelle II in Roth (2020b), S. 682.

Viertens weisen die ökonometrischen Ergebnisse, wie aus Tabelle 3 aus der Arbeit von Roth aus dem Jahr 2020[59] hervorgeht, auf die Bedeutung des Wachstums der immateriellen Kapitaldienstleistungen für das Wachstum der Arbeitsproduktivität auf der Makroebene hin. Roth verwendet einen Ansatz der

[59] Vgl. Roth (2020b).

„ländervergleichenden Wachstumsbilanzierung" für eine EU-16-Länderstichprobe im Zeitraum 2000 bis 2015. Sie basiert auf der um immaterielles Kapital erweiterten Modellspezifikation, wie sie zu Beginn dieses Beitrages vorgestellt wurde. Es zeigt sich, dass das Wachstum bei immateriellen Kapitaldienstleistungen den größten Teil des Wachstums der Arbeitsproduktivität erklären kann – bis zu 66 % beziehungsweise 46 %. Dies wird durch die Größe des Beta-Koeffizienten von 0,38 beziehungsweise 0,26 gezeigt. Ebenso signifikante, aber weniger ausgeprägte Ergebnisse finden sich auf der Meso- und der Mikroebene.[60]

Schlussfolgerungen

Der Beitrag führt zu dem Schluss, dass ein großer Teil des Produktivitätsrätsels gelöst werden kann, indem immaterielles Kapital in die Vermögensgrenze der Volkswirtschaftlichen Gesamtrechnungen einbezogen wird. Beim Produktivitätspuzzle handelt es sich also großenteils um einen Strukturwandel – von der Industrie- hin zur Wissensökonomik.

Exkurs: COVID-19-Krise und Arbeitsproduktivitätswachstum – ein Ausblick

Wie wird sich die gegenwärtige COVID-19-Krise auf das Wachstum der Arbeitsproduktivität auswirken? Um diese Frage zu beantworten, sollten wir eine kurzfristige von der mittel- bis langfristigen Perspektive unterscheiden.

Um die kurzfristigen Auswirkungen zu verstehen, ist es hilfreich, das Muster zu verstehen, das sich nach der Finanzkrise von 2008 abgezeichnet hat. In Abbildung 1.6 ihrer Publikation aus dem Jahr 2012 zeigt Mas die empirische Evidenz des Arbeitsproduktivitätswachstums für die EU und die USA für die Jahre 2007 bis 2010. Die Verfasserin verdeutlicht, dass die USA von 2007 bis 2010 einen tatsächlichen Anstieg des Arbeitsproduktivitätswachstums von 1,93 % auf 2,02 % verzeichneten, während für die EU-15 ein deutlicher Rückgang des Arbeitsproduktivitätswachstums von 1,41 % auf 0,07 % festzustellen war. Dieser Unterschied wird von Mas auf die je verschiedenen Arbeitsmarktregelungen in

[60] Vgl. Niebel et al. (2017), Marrocu et al. (2011)

den beiden Volkswirtschaften zurückgeführt.[61] Während die Wohlfahrtsstaaten der EU Kurzarbeitsprogramme intensiv genutzt haben, um die Gefahr großer Entlassungen nach der Finanzkrise einzudämmen, haben die USA von einer solchen Politik abgesehen.

Wie aus den Daten der Prognosen der Generaldirektion Wirtschaft und Finanzen (GDWF) der Europäischen Kommission vom Frühjahr 2020 hervorgeht, wird das Wachstum der Arbeitsproduktivität im Euroraum im Jahr 2020 um 3,2 Prozentpunkte zurückgehen, wobei in Deutschland ein Höchstwert von 5,6 Prozentpunkten erreicht werden wird.[62] Umgekehrt wird der Rückgang des Wachstums der US-Arbeitsproduktivität geringfügig sein und auf nur 0,2 Prozentpunkte geschätzt. Ähnlich wie nach der Finanzkrise im Jahr 2009 werden die Kurzarbeitsprogramme zur Dämpfung der Gefahr großer Entlassungen zu einem deutlichen Rückgang des Wachstums der Arbeitsproduktivität im Euroraum und in Deutschland gegenüber den USA führen. Aber wie groß ist die Auswirkung von COVID-19 aus historischer Sicht?

Eine aktuelle empirische Zeitreihenevidenz von Bergeaud et al. über den Zeitraum 1875 bis 2025 zeigt, dass die Auswirkungen auf das BIP-Wachstum zwar stärker sind als die der Finanzkrise im Jahr 2008, jedoch nur einen Bruchteil des Rückgangs ausmachen, der während der Weltwirtschaftskrise im Jahr 1929 zu verzeichnen war.[63] Darüber hinaus wird es im Jahr 2021 eine rasche Erholung und eine Rückkehr zum vorherigen Niveau von 2019 geben. Ein ähnlicher Rückgang der Investitionen im Jahr 2020 aufgrund der COVID-19-Krise mit einer starken Erholung im Jahr 2021 wird von der Generaldirektion Wirtschaft und Finanzen der Europäischen Kommission prognostiziert.[64] Ob diese auch für immaterielle Kapitalinvestitionen gilt, bleibt offen. Im Rahmen des GLOBALIN-

---

[61] Vgl. Mas (2012).

[62] Vgl. European Commission (2020a). Die GDWF benutzt Daten von Eurostat und berechnet das Wachstum der Arbeitsproduktivität im Jahr 2020 als gesamtjährliche prozentuale Veränderung des realen Bruttoinlandsprodukts pro Beschäftigten in Bezug auf das Gesamtjahr 2019. Im Gegensatz hierzu beruhen die Berechnungen von Schneider (2021) in diesem Sammelband auf Daten von Statista und auf der prozentualen Veränderung der Werte im zweiten Quartal 2020 in Bezug auf die Werte im vierten Quartal 2019 des realen Bruttoinlandsprodukts pro Gesamtheit der geleisteten Arbeitsstunden. Diese Unterschiede in Bezug auf Methodologie und Datenquelle führen zu den divergierenden Berechnungen des Arbeitsproduktivitätswachstums für das Fallbeispiel Deutschlands zu Zeiten der COVID-19-Krise.

[63] Vgl. Bergeaud et al. (2020).

[64] Vgl. European Commission (2020a).

TO-Projekts[65] hoffen die beteiligten Projektpartner, diese Frage mit einer passgenauen COVID-19-Umfrage zu immateriellen Kapitalinvestitionen – welche unter Zuhilfenahme einer repräsentativen Unternehmensstichprobe in sieben EU-Ländern durchgeführt wird – beantworten zu können.

Um die mittel- und langfristigen Auswirkungen zu verstehen, ist es hilfreich die politischen Maßnahmen, die zur Bewältigung der COVID-19-Krise ergriffen wurden, zu analysieren. Als Reaktion auf die Pandemie wurden auf der Ebene der Mitgliedstaaten unter den ausgewählten Kernländern des Euroraums historisch große Konjunkturpakete von bis zu 200 Mrd. Euro beschlossen.[66] Auf Bundesebene der EU beträgt die vereinbarte Haushaltskapazität insgesamt 750 Mrd. Euro.[67] Diese Finanzpolitik wird vom Pandemie-Notfallkaufprogramm (PEPP) der EZB mit einem Gesamtvolumen von 1.350 Mrd. Euro flankiert. Die Neuartigkeit des PEPP ist die Rolle, die die EZB als Kreditgeberin der letzten Instanz auf dem Markt für Staatsanleihen ohne Beschränkungen für Käufe in einzelnen Ländern übernimmt.[68] Ebenso historisch ist es für die Europäische Kommission, in ihrer Eigenschaft als multinationale Akteurin innerhalb ihres mehrjährigen Finanzrahmens 750 Mrd. Euro aufzunehmen. Dies ist höchstwahrscheinlich ein bedeutender Schritt, um eine stärkere Fiskalunion zu schaffen. Wie in einer Arbeit von Roth aus dem Jahr 2020 dargelegt, genießen die Präsidenten der EZB und der Europäischen Kommission angesichts der großen öffentlichen Unterstützung des Euro in den ersten zwei Jahrzehnten (1999 bis 2019) höchstwahrscheinlich die notwendige politische Legitimität, um diese entscheidenden Maßnahmen zu ergreifen.[69]

Werden diese Investitionspläne jedoch dazu beitragen, eine Erholung des Euroraums zu fördern? Aus obigen Argumenten lässt sich ableiten, dass diese Impulse dem Euroraum sicherlich helfen, sich kurzfristig zu erholen, insbesondere da es sich dieses Mal um ein dreifaches Impulsprogramm handelt: fiskalische Impulse auf der Ebene der Mitgliedstaaten und der EU, gepaart mit monetären Impulsen vonseiten der EZB.

Mittel- bis langfristig sind zwei Aspekte für die Erholung des Arbeitsproduktivitätswachstums relevant. Erstens muss der derzeitige Digitalisie-

---

[65] GLOBALINTO (2020).

[66] Vgl. Greive (2020)

[67] Vgl. European Commission (2020b).

[68] Vgl. Schnabel (2020)

[69] Vgl. Roth (2020c).

rungsschub mittels Investitionen aus den beschlossenen wirtschaftlichen Wiederaufbauplänen in die Digitalisierung und das notwendige ergänzende (unternehmerische und öffentliche) immaterielle Kapital unterstützt werden. Wenn die Mittel auf diese Weise verwendet werden, können wir erwarten, dass sich das Wachstum der Arbeitsproduktivität in der Zeit nach COVID-19 beschleunigt. Zweitens müssen die laufenden Investitionen in IKT und in immaterielle Vermögenswerte von wachstumsfördernden Reformen auf der Angebotsseite auf den Arbeits-, Produkt- und Dienstleistungsmärkten in den größeren Volkswirtschaften des Euroraums wie Italien flankiert werden. Dies sollte die notwendige Konvergenz der Lohnstückkosten gegenüber Deutschland bewirken.

Ein Post-COVID-19-Szenario lässt höchstwahrscheinlich einen deutlichen Anstieg des Arbeitsproduktivitätswachstums vermuten. Dies hängt jedoch davon ab, ob der gegenwärtig stattfindende Digitalisierungsschub mit Investitionen in die Digitalisierung sowie in das notwendige ergänzende (unternehmerische und öffentliche) immaterielle Kapital flankiert wird.

Literatur

Aghion, Philippe (2008): Higher Education and Innovation. In: Perspektiven der Wirtschaftspolitik, Vol. 9, S. 28–45.

Aghion, Philippe und Howitt, Peter (2006): Appropriate Growth Policy: A Unifying Framework. In: Journal of the European Economic Association, Vol. 4, S. 269–314.

Aghion, Philippe, Dewatripont, Matthias, Hoxby, Caroline, Mas-Colell, Andreu und Sapir, André (2007): Why Reform Europe's Universities? In: Bruegel Policy Brief (4), Bruegel, Brüssel.

Aghion, Philippe, Dewatripont, Matthias, Hoxby, Caroline, Mas-Colell, Andreu und Sapir, André (2008): Higher aspirations: An agenda for reforming European universities. In: Bruegel Blueprint Series (5), Bruegel, Brüssel.

Aghion, Philippe, Dewatripont, Matthias, Hoxby, Caroline, Mas-Colell, Andreu and Sapir, André (2010): The governance and performance of universities: evidence from Europe and the US. In: Economic Policy, Vol. 25 (1), S. 7–59.

Benhabib, Jess und Spiegel, Mark M. (1994): The role of human capital in economic development—evidence from aggregate cross-country data. In: Journal of Monetary Economics, Vol. 34 (2), S. 143–173.

Bergeaud, Antonin, Cette, Gilbert und Lecat, Rémy (2016): Productivity Trends in Advanced Countries between 1890 and 2012. In: Review of Income and Wealth, Vol. 62 (3), S. 420–444.

Bergeaud, Antonin, Cette, Gilbert und Lecat, Rémy (2020): Current and past recessions: a long-term perspective. EcoNotepad – post n° 159, Banque de France.

Bounfour, Antonin und Miyagawa, Tsutomu (2015): Intangibles, Market Failure and Innovation Growth, Springer Verlag, Heidelberg.

Brynjolfsson, Erik und Hitt, Lorin M. (2000): Beyond Computation: Information Technology, Organizational Transformation and Business Performance. In: Journal of Economic Perspectives, Vol. 14 (4), S. 23–48.

Brynjolfsson, Erik, Hitt, Lorin M. und Yang, Shinkyu (2002): Intangible Assets: Computers and Organizational Capital. In: Brookings Papers on Economic Activity, Vol. 2002 (1), S. 137–198.

Brynjolfsson, Erik, Rock, Daniel und Syverson, Chad (2019): Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In: Agrawal, Ajya, Gans, Joshua und Goldfarb, Avi (Hrsg.): The Economics of Artificial Intelligence: An Agenda, University of Chicago Press, Chicago, IL, S. 23–57.

Corrado, Carol, Hulten, Charles und Sichel, Daniel (2005): Measuring Capital and Technology: An Expanded Framework. In: Corrado, Carol, Haltiwanger, John und Sichel, Daniel (Hrsg.): Measuring Capital in the New Economy, University of Chicago Press, Chicago, IL, S. 11–46.

Corrado, Carol, Hulten, Charles und Sichel, Daniel (2009): Intangible capital and U.S. Economic growth. In: Review of Income and Wealth, Vol. 55 (3), S. 661–685.

Corrado, Carol, Haskel, Jonathan, Jona-Lasinio, Cecilia und Iommi, Massimiliano (2013): Innovation and intangible investment in Europe, Japan, and the United States, Oxford Review of Economic Policy 29, pp. 261–286.

Corrado, Carol, Haskel, Jonathan, Jona-Lasinio, Cecilia und Iommi, Massimiliano (2018): Intangible investment in the EU and US before and since the Great Recession and its contribution to productivity growth. In: Journal of Infrastructure, Policy and Development, Vol. 2, S. 11–36.

Cowen, Tyler (2011): The great stagnation: how America ate all the low-hanging fruit of modern history, got sick, and will (eventually) feel better, Dutton, New York.

De Grauwe, Paul (2015): Secular stagnation in the Eurozone. In: VoxEU, 30. Januar.

Draghi, Mario (2014): Unemployment in the Euro Area, Speech at the Annual Central Bank Symposium in Jackson Hole, Wyoming, 22. August.

Edquist, Harald (2011): Can Investment in Intangibles Explain the Swedish Productivity Boom in the 1990s? In: The Review of Income and Wealth, Vol. 57 (4), S. 658–682.

Europäische Kommission (2020a): European Economic Forecast, European Economy Institutional Paper 125, European Union, Luxemburg.

Europäische Kommission (2020b): Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the regions: The EU budget powering the recovery plan for Europe, COM/2020/442 final, Brüssel.

Fichtner, Ferdinand, Fratzscher, Marcel und Gornig, Martin (2014): Eine Investitionsagenda für Europa, DIW Wochenbericht, Nr. 27, S. 631–635.

Fratzscher, Marcel (2014): Die Deutschland-Illusion. Warum wir unsere Wirtschaft überschätzen und Europa brauchen, Carl Hanser Verlag, München.

Fukao, Kyoju, Miyagawa, Tsutomu, Mukai, Kentaro, Shinoda, Yukio und Tonogi, Konomi (2009): Intangible Investment in Japan: Measurement and Contribution to Economic Growth. In: Review of Income and Wealth, Vol. 55 (3), S. 717–736.

GLOBALINTO (2020): Capturing the value of intangible assets in micro data to promote the EU's growth and competitiveness. Online abrufbar unter https://globalinto.eu, zuletzt aufgerufen am 2.6.2021.

Gordon, Robert J. (2018): Declining American Growth despite ongoing innovation. In: Explorations in Economic History, Vol. 69 (C), S. 1–12.

Greive, Martin (2020): Bund macht fast 220 Milliarden Euro neue Schulden. In: Handelsblatt, 15. Juni.

Haskel, Jonathan und Westlake, Stian (2018a): Capitalism without Capital, Princeton University Press, Princeton.

Haskel, Jonathan und Westlake, Stian (2018b): Productivity and secular stagnation in the intangible economy. In: VoxEU, 31. Mai.

Haskel, Jonathan (2020): Monetary Policy in the intangible economy. Rede für die Bank of England, gehalten am 11. Februar in der Universität Nottingham.

Heil, Mark (2018): Finance and Productivity: A Literature Review. In: Journal of Economic Surveys, Vol. 32 (5), S. 1355–1383.

INNODRIVE (2011): INNODRIVE Intangibles Database.

Krugman, Paul (1994): Competitiveness: A Dangerous Obsession. In: Foreign Affairs, Vol. 73 (2), S. 28–44.

Krugman, Paul (2014): Depressions are different. In: Solow, Robert M. und Murray, Janice (Hrsg.): Economics for the curious, Palgrave Macmillan, New York, S. 7–18.

Marrano, Mauro G., Haskel, Jonathan und Wallis, Gavon (2009): What happened to the knowledge economy? ICT, intangible investment, and Britain's productivity record revisited. In: Review of Income and Wealth, Vol. 55 (3), S. 686–716.

Marrocu, Emanuela, Paci, Raffaele und Pontis, Marco (2011): Intangible Capital and firms' productivity. In: Industrial and Corporate Change, Vol. 21 (2), S. 377–402.

Mas, Matilde (2012): Productivity in the Advanced Countries: From Expansion to Crisis. In: Mas, Matilde und Stehrer, Robert (Hrsg.): Industrial Productivity in Europe, Edward Elgar, Cheltenham, S. 11–34.

Nakamura, Leonard (2010): Intangible Assets and National Income Accounting. In: Review of Income and Wealth, Vol. 56 (s1), S. 135–155.

Nakamura, Leonard (2019): Economic Growth--Too Slow or Too Fast? Measuring the Value of Intangibles and AI. Vortrag auf der 15. Weltkonferenz zum Intellectual Capital for Communities am 11. und 12. Juli, UNESCO, Paris.

Niebel, Thomas, O'Mahony, Mary und Saam, Marianne (2017): The Contribution of Intangible Assets to Sectoral Productivity Growth in the EU. In: The Review of Income and Wealth, Vol. 63 (s1), S. 49–67.

OECD (2015): The future of productivity, Paris, OECD.

Oulton, Nicholas (2018): Productivity and the Great Recession. In: Intereconomics, Vol. 53 (2), S. 63–68.

Posen, Adam S. und Zettelmeyer, Jeromin (2019): Facing Up to Low Productivity Growth: Introduction. In: Posen, Adam S. und Zettelmeyer, Jeromin (Hrsg.): Facing Up to Low Productivity Growth, Columbia University Press, New York, S. 1–11.

Remes, Jaana, Mischke, Jan und Krishnan, Mekala (2018): Solving the productivity puzzle: The role of demand and the promise of digitization. In: International Productivity Monitor, Nr. 35, S. 28–51.

Ross, Jenna (2020): Intangible Assets: A Hidden but Crucial Driver of Company Value. Online abrufbar unter: https://www.visualcapitalist.com/intangible-assets-driver-company-value, zuletzt aufgerufen am 2.6.2021.

Roth, Felix (2019): Intangible Capital and Labour Productivity Growth: A Literature Review, Hamburg Discussion Papers in International Economics, Nr. 4, Hamburg.

Roth, Felix (2020a): The Productivity Puzzle – A Critical Assessment. Habilitationsvortrag vor dem Habilitationskomitee der Fakultät für Wirtschafts- und Sozialwissenschaften der Universität Hamburg am 29. Juni.

Roth, Felix (2020b): Revisiting Intangible Capital and Labour Productivity Growth, 2000–2015: Accounting for the Crisis and Economic Recovery in the EU. In: Journal of Intellectual Capital, Vol. 21 (5), S. 671–690.

Roth, Felix (2020c): Economic Recovery Strengthens Public Support for the Euro. In: JCMS-Blog, 9. Januar.

Roth, Felix (2021): The Productivity Puzzle – A Critical Assessment and an Outlook on the COVID-19-Crisis. In: Roth, Felix: Intangible Capital and Growth – Essays on Labor Productivity, Monetary Economics, and the Political Economy, Vol. 1, Springer, New York, Kapitel 1.

Roth, Felix und Thum, Anna-Elisabeth (2013): Intangible Capital and Labour Productivity Growth. In: Review of Income and Wealth, Vol. 59 (3), S. 486–508.

Sapir, André, Aghion, Philippe, Bertola, Giuseppe, Hellwig, Martin, Pisani-Ferry, Jean, Rosati, Dariusz, Viñals, José und Wallace, Helen (2004): An Agenda for a Growing Europe: The Sapir Report, Oxford University Press, Oxford.

Schnabel, Isabel (2020): The ECB's policy in the COVID-19 crisis – a medium-term perspective, Europäische Zentralbank, Rede am 10.6.2020, Frankfurt am Main.

Schneider, Henrique (2021): Zwei Rätsel der Produktivität – eine kritische Bewertung. In: Straubhaar, Thomas (Hrsg.): Neuvermessung der Datenökonomie, Hamburg University Press, Hamburg, S. 84–99.

Summers, Lawrence H. (2014): Reflections on the „New Secular Stagnation Hypothesis". In: VoxEU, 15. August.

Summers, Lawrence H. (2015): Demand Side Secular Stagnation. In: American Economic Review, Vol. 105 (5), S. 60–65.

Syverson, Chad (2017): Challenges to Mismeasurement Explanations for the US Productivity Slowdown. In: Journal of Economic Perspectives, Vol. 31 (2), S. 165–86.

Teulings, Coen und Baldwin, Richard (2014): Secular Stagnation: Facts, Causes and Cures. In: VoxEU, 15. August.

Van Ark, Bart, O'Mahony, Mary und Timmer, Marcel P. (2008): The Productivity Gap between Europe and the United States: Trends and Causes. In: Journal of Economic Perspectives, Vol. 22 (1), S. 25–44.

Van Ark, Bart (2016): Europe's Productivity Slowdown Revisited. A Comparative Perspective to the United States, In: Askenazy, Philippe, Bellmann, Lutz, Bryson, Alex und Moreno Galbis, Eva (Hrsg.): Productivity Puzzles Across Europe. Oxford University Press, Oxford, S. 26–48.

Van Ark, Bart und Jäger, Kirsten (2017): Recent Trends in Europe's Output and Productivity Growth Performance at the Sector Level, 2002–2015. In: International Productivity Monitor, Vol. 33, S. 8–23.

Van Ark, Bart und O'Mahony, Mary (2016): Productivity Growth in Europe before and since the 2008/2009 economic and financial crisis, In: Jorgensen, Dale, Fukao, Kyoji und Timmer, Marcel P. (Hrsg.): The World Economy. Growth Stagnation? Cambridge University Press, Cambridge, S. 111–152.

Van Ark, Bart, de Vries, Klaas und Jäger, Kirsten (2018): Is Europe's Productivity Glass Half Full or Half Empty? In: Intereconomics, Vol. 53, S. 53–58.

Zwei Rätsel der Produktivität – eine empirische Beobachtung

Henrique Schneider

Einleitung

Neben dem „Rätsel der Produktivität" in seiner üblichen Form – entwickelte, investitionsintensive und „digitale" Wirtschaften weisen einen nur geringen Anstieg der Produktivität aus – hätte man im Jahr 2020 ein zweites Rätsel erkennen können: Wie kann es sein, dass gerade in einer Zeit der Krise oder Rezession die Arbeitsproduktivität steigt?

Während des ersten Halbjahres 2020 fanden in verschiedenen europäischen Ländern sogenannte Lockdowns statt. Es waren Maßnahmen zur Eindämmung der COVID-19-Pandemie. Einzelne wirtschaftliche Aktivitäten wie etwa der Handel oder die Gastronomie konnten gar nicht mehr ausgeführt werden. Andere Sektoren, so zum Beispiel die Industrie oder die Finanzdienstleistungen, konnten nur unter Einhaltung von Schutzkonzepten weiter funktionieren. Heim- und andere Modalitäten der Fernarbeit wurden empfohlen und praktiziert. Die Wertschöpfung brach ein. Mit dem Anstieg der Arbeitslosigkeit und der Beanspruchung von Kurzarbeit nahm die Zahl der geleisteten und bezahlten Arbeitsstunden ebenfalls ab. Gerade in diesem Zusammenhang stieg die Arbeitsproduktivität aber markant an – gemäß der Statistik.

Dieser Beitrag geht dem Rückgang des Produktivitätsfortschritts aus empirischer Perspektive nach und findet in der Steigerung der Produktivität im ersten Halbjahr 2020 eine weitere Facette, die zu beobachten ist: die Multifaktorproduktivität.

Nach einer kurzen Rekapitulation der Ausgangslage folgt eine empirische Darstellung der Produktivitätsentwicklung im ersten Halbjahr 2020. Daran schließt sich eine ebenso empirisch begründete Diskussion der Multifaktorproduktivität an. Abschließend werden Schlussfolgerungen gezogen und Desiderate für weitere die Forschung geortet.

Ausgangslage: Forschung und Empirie

Allgemein-abstrakt ist Produktivität das Verhältnis von Output zu Input, wobei vor allem die Entwicklung dieses Verhältnisses im Lauf der Zeit von Interesse ist. In der ökonomischen Theorie, vor allem in der Makroökonomie haben sich drei Produktivitätsmaße etabliert: die Arbeitsproduktivität, die Kapitalproduktivität und die Multifaktorproduktivität (auch totale Faktorproduktivität genannt).

Die Arbeitsproduktivität entspricht der realen Wertschöpfung pro geleisteter Arbeitsstunde. Volkswirtschaftlich kann man die Veränderungsrate der Arbeitsproduktivität definieren als die Differenz zwischen der Entwicklungsrate des Bruttoinlandsprodukts (BIP) und der Veränderung des Arbeitsinputs.[1]

Die Kapitalproduktivität misst die durch die Kapitalleistungen realisierte Wertschöpfung. Das Wachstum der Kapitalproduktivität entspricht der Differenz zwischen der Entwicklung des Bruttoinlandsprodukts und der Veränderung des Kapitalinputs.[2]

Sowohl mit der Arbeitsproduktivität als auch mit der Kapitalproduktivität werden nur bestimmte Elemente der Produktivität gemessen. Man spricht deshalb von Teilmaßen. Die Produktivität hängt aber von sämtlichen Produktionsfaktoren und ihrer Kombination ab. Es ist deshalb schwierig, das tatsächlich für die Produktivität verantwortliche Element zu isolieren. Mit der Multifaktorproduktivität lässt sich die volkswirtschaftliche Effizienz umfassender messen. Sie entspricht der Bruttowertschöpfung zu Vorjahrespreisen pro kombinierter Faktoreinheit (Kapital, Arbeit usw.). Bei dieser Messung werden gleichzeitig Arbeit und Kapital berücksichtigt, womit der Beitrag dieser beiden Produktionsfaktoren zum Wirtschaftswachstum als ein Bündel betrachtet wird. Die Entwicklungsrate der Multifaktorproduktivität entspricht der Differenz zwischen der Veränderung der Bruttowertschöpfung und der Veränderung des gesamten Faktorinputs.[3]

Die Arbeitsproduktivität ist das am breiteste verwendete volkswirtschaftliche Produktivitätsmaß. Sie beeinflusst Tarifverhandlungen oder Erwägungen zum Mindestlohn und findet deshalb in der Presse und in der Politik erhöhte Aufmerksamkeit.[4] Die Entwicklung der Arbeitsproduktivität ist in Verlangsa-

---

[1] OECD (2015).

[2] OECD (2015).

[3] Schneider (2020).

[4] Burda (2018).

mung begriffen. Das gilt für Deutschland, für Europa und für praktisch alle entwickelten Wirtschaften, etwa für die Mitglieder der Organisation für ökonomische Zusammenarbeit und Entwicklung (OECD).[5]

Eine Übersicht über diese Entwicklung bietet Abbildung 1. Sie zeigt die Veränderung der Arbeitsproduktivität im Vorjahresvergleich in Prozent für die 19 Mitglieder der Eurozone, die 28 Länder der EU, die gesamte OECD-Gruppe sowie für Deutschland, Frankreich und die USA.[6]

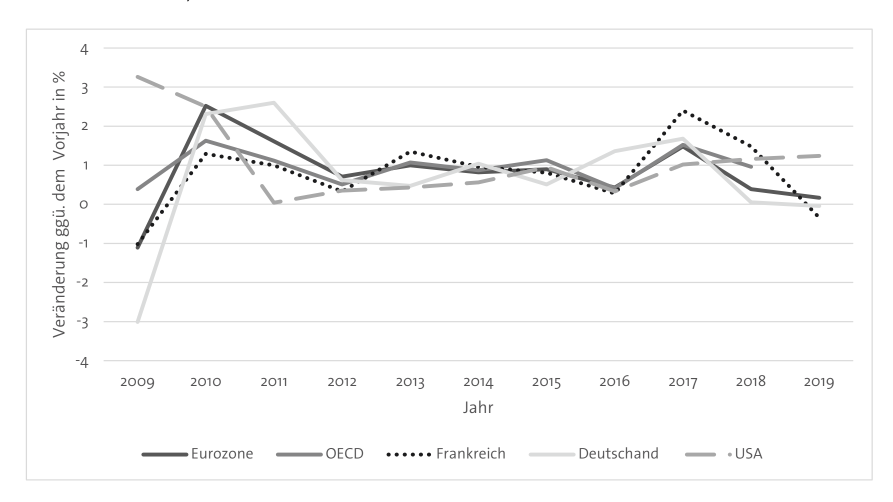


Abb. 1: Entwicklung der Arbeitsproduktivität in Prozent im Vergleich mit dem Vorjahr
Quelle: OECD (2020a).

Die hier betrachtete Zeitreihe fängt mit dem Jahr 2009 an. Damals waren die Veränderungsraten der Produktivität gering oder sogar negativ. In den unmittelbaren Folgejahren nahmen sie zu. Die naheliegende Interpretation ist, dass die Wertschöpfung pro Arbeitsstunde während der Großen Rezession 2007 bis 2009 abnahm und dann in der Erholungsphase bis zum Jahr 2011 wieder zunahm. In den Jahren 2012 bis 2016 pendelte sich das Wachstum der Arbeitsproduktivität auf um etwa 1 % pro Jahr ein. In den Jahren 2016 und 2017 stieg diese Veränderung an, doch in den Jahren 2018 und 2019 fiel sie stark – bis etwa zur Nullmarke.

---

5 Kuntze/Mai (2020).

6 OECD (2020a). Die Berechnung erfolgt mittels der oben genannten Technik als Output, gemessen in Bruttoinlandsprodukt, und Input, gemessen an der Zahl der geleisteten Arbeitsstunden.

An dieser Stelle gilt es, zweierlei zu beachten. Erstens legen diese Zahlen nahe, dass Deutschland die Entwicklung der Produktivität in den Vergleichsgruppen jeweils um etwa ein Jahr vorwegnimmt. Das hat vermutlich mit der Stellung der deutschen Wirtschaft als „Konjunkturmotor“ zu tun.[7] Zweitens entwickelt sich die Produktivität in einem schmalen Rahmen. Gerade diese Entwicklung gilt als einer der Belege für das „Rätsel der Produktivität“.

Es bestehen verschiedene Erklärungen für dieses „Rätsel“, wobei sich kein Forschungskonsens erkennen lässt. Auf der einen Seite versuchen einige, die Verlangsamung der Entwicklung der Arbeitsproduktivität mit Rückgriff auf die Digitalisierung, auf „Baumol’s disease“, auf Regulierung oder auf den Zusammenbruch des Innovations-Diffusions-Mechanismus zu erklären. Während ein Teil der Forschung sagt, aus der Digitalisierung folgten keine wesentlichen Impulse für die Erhöhung der Produktivität,[8] meint der andere Teil, die Digitalisierung befinde sich erst in einer Aufbauphase. Wenn diese Phase einmal abgeschlossen sei, entwickelten sich die Produktivitätspotenziale.[9] „Baumol’s disease“ bezeichnet die Problematik der schlechten Rationalisierbarkeit von Dienstleistungen im Vergleich mit den Outputs anderer Sektoren. Als Bereitsteller von (Über-)Kapazitäten hat der Dienstleistungssektor von vornherein höhere Fixkosten und ist daher grundsätzlich weniger produktiv – im Sinne der üblichen Messtechnik.[10]

Ein anderer Ansatz findet die Erklärung für das Rätsel in der zunehmenden Regulierung, welche auch Handelshemmnisse einführt. Regulierung wirkt sich dabei als nicht rationalisierbarer Fixkostenblock aus, was wiederum zu hoher Kostenremanenz führt.[11] Der Zusammenbruch des Innovations-Diffusions-Mechanismus sagt wiederum, dass derzeit nur wenige Unternehmen produktivitätssteigernde Innovationen entwickeln. Zudem bleiben diese Innovationen proprietär und werden von anderen Unternehmen weder nachgeahmt noch integriert.[12] Damit entwickelt sich eine zunehmende Disparität zwischen innovativen und produktiven Unternehmen und solchen, die es nicht sind.

---

[7] Karremans (2020).

[8] Gordon (2016).

[9] Van Ark (2016).

[10] Baumol/Bowen (1966).

[11] Lemieux (2018).

[12] OECD (2015).

Auf der anderen Seite zweifelt die sogenannte Mismeasurement-Hypothese an der statistischen Grundlage der Messung der Produktivität. Gemäß dieser Hypothese müsste das Bruttoinlandsprodukt um „Intangibles“ erweitert werden. Diese beinhalten etwa Güter, die gratis sind – so die Suche im Internet und viele digitale Applikationen –, Qualitätsverbesserungen oder auch das geistige Eigentum. Daraus folgt, dass die Messtechniken zur Erfassung sowohl der Wertschöpfung als auch der Produktivität einer Aktualisierung bedürfen. Arbeiten an einer entsprechenden Erweiterung der Konzepte und Techniken werden unter anderem in dieser Publikation vorgestellt. Weitere Arbeiten sind etwa die CHS-Referenzrahmen[13] und die Arbeiten der KLEMS-Initiative.[14]

Eine Überprüfung dieser Erklärungen anhand der zur Verfügung stehenden Daten zeigt kein konklusives Bild. Die oben angesprochenen Ansätze werden durch die Daten teilweise plausibilisiert und teilweise widerlegt.[15] Aber auch die „Mismeasurement“-Hypothese ist umstritten. Auch die unterschiedlichen statistischen Erweiterungen des Bruttoinlandsprodukts haben bisher nicht zu einer nennenswert höheren Produktivität geführt.[16]

Eine weitere Lücke ist, dass viel weniger über die Kapital- und Multifaktorproduktivität geforscht wird als über die Arbeitsproduktivität.[17] Diese beide Aspekte werden meist als Erklärungsfaktoren der Arbeitsproduktivität behandelt. Zwar werden mittlerweile verschiedene Formen des Kapitalaufbaus unterschieden – man ist der Digitalisierung entgegengekommen und unterschiedet zwischen ICT- und Nicht-ICT-Kapital. ICT steht für Informatik und Telekommunikation beziehungsweise „information and communications technology“.[18] Doch schon etwa bei der Multifaktorproduktivität gibt es kaum Fortschritte in der Forschung.

Der vorliegende Beitrag versucht, diese Forschungslücke zu schließen, indem er die Änderung der Multifaktorproduktivität als mögliche Erklärung für den Anstieg der Arbeitsproduktivität im ersten Halbjahr des Jahres 2020 heranzieht. Dies geht mit einer Erweiterung des Konzepts einher.

---

13 Corrado/Hulten/Sichel (2009), Corrado/Hulten (2010).

14 Jorgenson (2017).

15 Schneider (2020).

16 Syverson (2017).

17 Ahmed/Bhatti (2020).

18 Cardona et al. (2013).

Produktivität während des COVID-19-Lockdowns

Während des ersten Halbjahrs 2020, der Zeit des sogenannten Lockdowns, ist die Produktivität der Arbeit sprunghaft angestiegen. Abbildung 2 zeigt, dass in den zum Vergleich herangezogenen Ländern die Produktivität während des ersten Halbjahres 2020 höher ausfiel als im Durchschnitt der Dekade 2009 bis 2019. In Deutschland kletterte die Wachstumsrate von 1 % auf mehr als 3 %, in Frankreich auf fast 6 %, und auch in den USA kam zu einer Erhöhung von 0,5 % auf 1,7 %.[19]

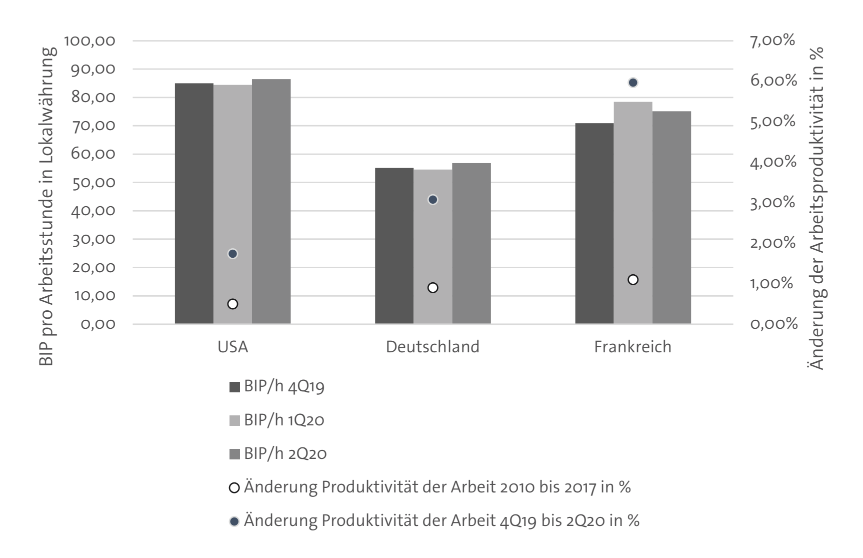


Abb. 2: Entwicklung der Arbeitsproduktivität in Prozent im ersten Halbjahr 2020
Quelle: Schneider (2020).

Freilich ist im Umgang mit solchen Zahlen und ihren Veränderungen Vorsicht geboten. Sie beinhalten einen Basiseffekt. Zwar ist die Wertschöpfung in den betroffenen Ländern eingebrochen, doch die Zahl der geleisteten Arbeitsstunden ist noch stärker zurückgegangen. Wenn der Nenner stärker abnimmt als der Zähler, ist die Änderungsrate positiv, nimmt die Produktivität also zu. Selbst hier ist noch zu berücksichtigen, dass einige dieser Zahlen provisorische Schät-

[19] Schneider (2020).

zungen sind beziehungsweise zum Zeitpunkt der Entstehung der Studie, aus der sie stammen, waren.[20]

Trotzdem ist die Arbeitsproduktivität während des Lockdowns angesprungen. Interessant in diesem Zusammenhang ist nicht die genaue Bezifferung des Anstiegs, sondern das Spektrum der Erklärungen dafür. Die relevantesten Gruppen von Erklärungen sind:

1. *Deregulierung und Digitalisierung.* Zumindest in der kurzen Frist haben Deregulierung und Digitalisierung die Produktivität erhöht. Während der COVID-19-Pandemie lockerten die meisten Wirtschaften ihre starren Arbeitsgesetze und ließen vermehrt Fern- und Telearbeit zu. Gleichzeitig öffneten sich die Unternehmen dafür, vermehrt Technologien in die entsprechenden Arbeitsprozesse ihrer Mitarbeitenden zu integrieren. Es stellt sich nun die Frage, ob die Regulatoren und der private Sektor bereit sind, die Deregulierung und Digitalisierung fortzusetzen. Falls ja, könnte die Produktivität weiter steigen – höchstwahrscheinlich nicht mit der gleichen Geschwindigkeit wie während dieses kurzen hier betrachteten Abschnitts, aber zumindest mit einer höheren Rate als im Durschnitt der Jahre von 2010 bis 2018.
2. *Steuerung und Verbesserung betrieblicher Prozesse.* Die Kombination aus Rationalisierung von Prozessen, durchdachter Koordination und Konzentration auf selbstgesteuerte Arbeit ist eine weitere Erklärung für den Produktivitätsanstiegs. In diesem Ansatz geht es nicht um exogen gegebene Technologie oder Regulierung, sondern um endogen in den Betrieben erzielte Steuerungs- und Verbesserungsgewinne. Diese wurden etwa mit Prozessorganisation, „lean management“, „Management by Objectivess“ („MbO“) oder der Übertragung einer gewissen industriellen Logik auf Dienstleistungen (Schichten, Takt, Ergebniskontrolle, Kapazitätsminimierung) erzielt. So gewonnene Produktivitätssteigerungen können auch unabhängig von Digitalisierung und Deregulierung umgesetzt werden.
3. *„Mismeasurement“ und Externalitäten.* Auch in dieser Situation ließe sich die „Mismeasurement“-Hypothese formulieren: Der Produktivitätsschub kam

[20] Der Beitrag von Felix Roth in diesem Band kommt zu einem anderen Ergebnis. Roth stellt heraus, dass die Produktivität pro erwerbstätiger Person im gleichen Bezugsjahr abnimmt. Dieser Unterschied geht auf die unterschiedlichen Messtechniken zurück. Roth stützt sich auf Berechnungen der Europäischen Kommission, welche die Produktivität als Wertschöpfung pro erwerbstätiger Person berechnet; dieser Beitrag nimmt das OECD-Maß, nämlich die Wertschöpfung pro geleisteter Arbeitsstunde. Zur Plausibilisierung: Wenn eine erwerbstätige Person in Kurzarbeit ist, bleibt sie erwerbstätig und damit Teil des Nenners in der Berechnung; da sie aber nicht oder weniger arbeitet, entfallen geleistete Arbeitsstunden, was wiederum den Nenner der Berechnung kleiner macht.

nur zustande, weil die ökonomischen Messgrößen die mit den gesundheitspolitischen Maßnahmen einhergehenden Externalitäten nicht berücksichtigen. In diesem Fall handelt es sich bei den Externalitäten um die höhere Wahrscheinlichkeit von Depressionen oder anderen psychologischen Auswirkungen auf die Menschen, die nicht oder weniger arbeiten; um mehr Stress für Menschen, die mit erhöhter Flexibilität oder aus der Ferne arbeiten; oder um den Verlust sozialer Kontakte am Arbeitsplatz. Einige dieser Externalitäten bleiben auch dann bestehen, wenn die Menschen zu ihren Arbeitsroutinen zurückkehren. Was auf den ersten Blick wie ein Produktivitätsschub aussieht, ist gemäß diesem Ansatz eine Fehlanzeige.

Insgesamt und aus volkswirtschaftlicher Perspektive kann der signifikante Anstieg der Produktivität im ersten Halbjahr 2020 zumindest als Indikator für entsprechende noch nicht umgesetzte Produktivitätspotenziale gelten. Der Indikator könnte also bedeuten, dass Arbeitszeit vorhanden ist, die bezahlt, aber nicht ausgelastet wird, oder dass Arbeitszeit mittels des Einsatzes von Technologie oder Flexibilität oder einfach durch besseres Management zu mehr Output führen kann.

Die erste und die zweite Gruppe von Erklärungen, so unterschiedlich sie sind, führen die Erhöhung der Arbeitsproduktivität auf die Disposition von Arbeit beziehungsweise auf die Kombination von Arbeit und Kapital zurück. Man kann sie auch als Freisetzungsproduktivitätsgewinn bezeichnen. Gemäß beiden Gruppen von Ansätzen ist der Anstieg der Arbeitsproduktivität auf die Multifaktorproduktivität zurückzuführen. Die dritte Gruppe von Erklärungen wählt einen anderen Weg.. Gemäß ihr ist die Arbeitsproduktivität schlicht ein schlechtes Maß oder sie kann nicht alle Externalitäten einbeziehen, die der Lockdown mit sich brachte. Dieser Vorwurf, so ernst er zu nehmen ist, steht nicht im Widerspruch zu den beiden ersten Gruppen, sondern ergänzt sie. Darauf wird später zurückzukommen sein.

Multifaktorproduktivität

Die Multifaktorproduktivität wird als Residualgröße berechnet; entsprechend wurde sie lange als Produktivitätsgewinn, der sich nicht eindeutig auf Arbeit oder Kapital zurückführen lässt, verstanden.[21] Doch im Laufe der Überprüfung

---

[21] Cantner et al. (2007).

der „Mismeasurement"-Hypothese stellten verschiedene Gruppen von Forschenden eine leichte, aber immerhin positive Korrelation zwischen der Kapitalintensivierung allgemein und der Multifaktorproduktivität fest. Seither versuchen sie, die Multifaktorproduktivität als Technologieproduktivität zu lesen. Sie interpretieren die Multifaktorproduktivität auch als „kostenlose" Innovation oder als Dispersion von Wissen seitens der Innovatoren auf andere Akteure.[22]

Das ist eine mögliche, aber zu eng gefasste Interpretation der Multifaktorproduktivität. Selbst wenn ein Zusammenhang zwischen der Kapitalintensivierung und der Multifaktorproduktivität besteht, heißt dies noch nicht, dass Änderungen der Multifaktorproduktivität allein auf Kapitalintensivierung zurückzuführen sind. Als Maß für die Gesamtproduktivität kann die Multifaktorproduktivität alles inkludieren, was in nicht „perfekten" Märkten passiert: Neben dem technologischen Fortschritt gehören dazu unter anderem auch Innovationseffekte, Änderungen von Organisationsformen und Institutionen, Messfehler, Skaleneffekte und Externalitäten.[23]

Eine weitere Interpretation der Multifaktorproduktivität ist, sie als Maß für die Qualität des Managements zu sehen. Management wird dabei als Unternehmensführung verstanden, das heißt als Auswahl und Einsatz von Arbeit, als Entscheid über Kapitalinvestitionen und als Disposition von Arbeit und Kapital. Die beiden im vorherigen Abschnitt genannten Gruppen von Erklärungen scheinen eher dieses Verständnis der Multifaktorproduktivität zu plausibilisieren. Als die Daten zur Produktivitätsentwicklung im ersten Halbjahr 2020 ausgewertet wurden, wurden die Ergebnisse im Rahmen eines Panels mit Unternehmerinnen und Unternehmern validiert. Die Teilnehmenden zeigten sich über den Produktivitätszuwachs während des Lockdowns nicht erstaunt. Auch sie führten die Steigerung auf Management-Maßnahmen zurück – so wie die erste und zweite Gruppe von Erklärungen es taten. Die Teilnehmenden blieben jedoch vorsichtig. Sie äußerten die Vermutung, der erhöhte Produktivitätsstandard lasse sich nicht ewig halten. Die meisten Teilnehmenden waren auch der Überzeugung, dass verschiedene Dinge, etwa Externalitäten oder die Vorläufigkeit der Maßnahmen, mit aktuellen Methoden gar nicht gemessen werden könnten. Also doch „Mismeasurement".[24]

---

[22] Van Ark et al. (2009); Corrado et al. (2009).

[23] Hulten (2001).

[24] Schneider (2020).

Trotz möglicher Verkürzungen eröffnet die Interpretation der Multifaktorproduktivität als Qualität des Managements gewisse Perspektiven. Sie erlaubt beispielsweise eine Verknüpfung mit der Effizienzdiskussion in der Betriebswirtschaftslehre. Die Steuerung und Verbesserung der Produktivität wird dort als eine der wichtigsten Aufgaben des Managements verstanden. Entsprechend wurden dafür verschiedene Erfassungs- und Steuerungsinstrumente entwickelt, zum Beispiel die „Balanced Scorecard" oder das „Benchmarking." Eine so verstandene Multifaktorproduktivität würde als volkswirtschaftliche Kennzahl die betriebswirtschaftlichen Fortschritte einer Wirtschaft erfassen und damit unter anderem auch die Diffusion von Wissen, Know-how, Technologie einbeziehen.[25] Als Produktivitätsgesamtmaß kann die so verstandene Multifaktorproduktivität auch die „Intangibles" umfassen.

Vorsicht ist geboten. Auch in dieser Interpretation wäre die Multifaktorproduktivität zwar ein Indikator für die Qualität des Managements, was wiederum die Adaption von Technologie, die Dispersion von Wissen und Knowhow, die Innovation und anderes beinhaltet, aber sie wäre dies nur aus höherer Warte tun. Eine genaue Zerlegung der Bestimmungsfaktoren der Qualität des Managements ist mit einer Residualgröße nicht möglich. Dennoch: Was zeigt eine empirische Untersuchung der Multifaktorproduktivität? Werden die oben angesprochenen Perspektiven tatsächlich eröffnet?

Eine empirische Anzeige

Empirisch gehört die Multifaktorproduktivität zu den am wenigsten recherchierten Gebieten der Ökonomik. Auch die entsprechenden statistischen Ämter sind meist zurückhaltend in der Publikation von Zahlen dazu. Abbildung 3 zeigt die Entwicklung der Multifaktorproduktivität in Deutschland, Frankreich und den USA in den letzten zehn Jahren. Es handelt sich jeweils um die prozentuale Änderung im Vergleich mit dem Vorjahr.

Gerade im Vergleich mit der Entwicklung der Arbeitsproduktivität fällt die leicht höhere Volatilität der Multifaktorproduktivität auf. In Deutschland entwickelt sich die Multifaktorproduktivität leicht stärker als die Arbeitsproduktivität. Das ist in den USA und Frankreich umgekehrt.

---

[25] Erken et al. (2018).

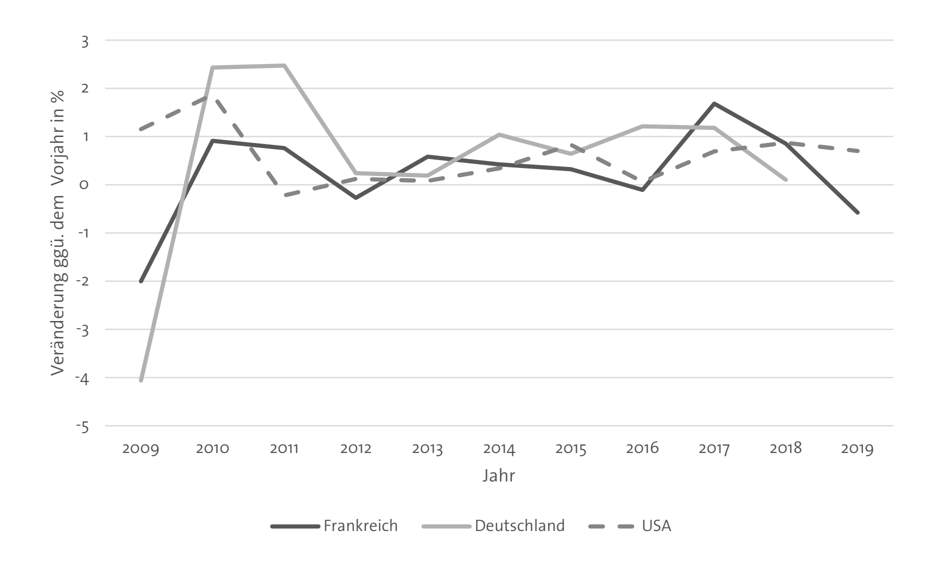


Abb. 3: Entwicklung der Multifaktorproduktivität in Prozent im Jahresvergleich
Quelle: OECD (2020b).

Die OECD nimmt in ihren „Productivity Statistics 2019" eine genauere Untersuchung der Entwicklung der Multifaktorproduktivität vor.[26] Abbildung 4 stellt das Ergebnis dieser Analyse dar. Dort werden Deutschland mit Frankreich und den USA über eine lange Periode von 1995 bis 2017 verglichen. Diese Periode wird in zwei kürzere eingeteilt, 2001 bis 2007 und 2010 bis 2017. Über die lange Periode weisen alle Länder ein durchschnittliches Wachstum der Multifaktorproduktivität von weniger als 1 % pro Jahr auf, wobei die USA den höchsten Durchschnitt erzielt, gefolgt von Deutschland und Frankreich. Konzentriert man sich allerdings auf die kürzeren Perioden, so stellt man fest, dass die USA und Frankreich eine Verlangsamung des Wachstums der Multifaktorproduktivität verzeichnen, während sich dieses in Deutschland stetig erhöht. Die deutsche Wirtschaft steigert also die Produktivität der gesamten Wirtschaft; ihr Durchschnitt für die Periode 2010 bis 2017 lag bei beinahe 1 % und damit deutlich höher als in den USA und Frankreich.

[26] OECD (2019).

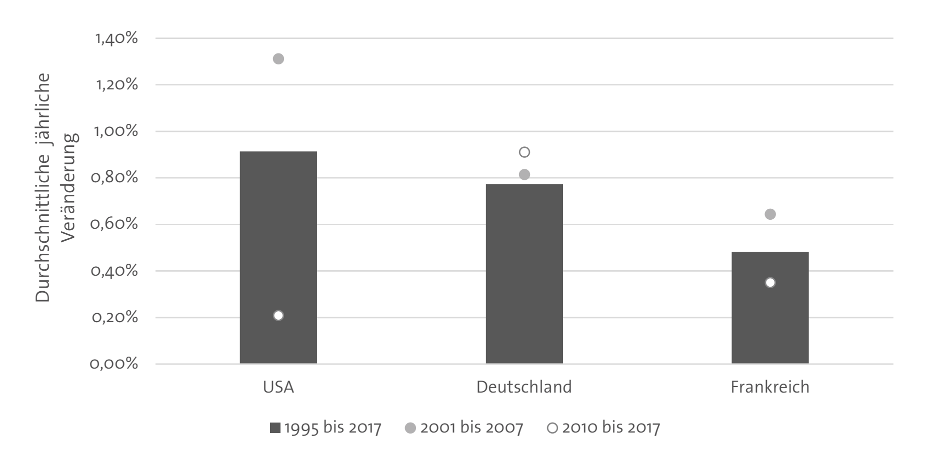


Abb. 4: Entwicklung der Multifaktorproduktivität in einer langen Periode und unterteilt in zwei kürzere
Quelle: OECD (2019).

Die OECD[27] geht einen Schritt weiter und versucht herauszufinden, welchen Effekt die Multifaktorproduktivität auf die Arbeitsproduktivität hat. Die Arbeitsproduktivität kann analytisch auch als eine Funktion von Kapitalintensivierung und Multifaktorproduktivität bestimmt werden.[28] Dabei wird vermutet: Je höher die Kapitalintensivierung, desto höher die Arbeitsproduktivität; je höher die Multifaktorproduktivität, desto höher die Arbeitsproduktivität; und je höher das Produkt aus Kapitalintensivierung und Multifaktorproduktivität, desto höher die Arbeitsproduktivität.

Die OECD unterscheidet dabei zwei Formen der Kapitalintensivierung: jene des ICT-Kapitals – das ist der Kapitalbestand in Informations- und Kommunikationstechnik – und jene des Nicht-ICT-Kapitals.

Abbildung 5 zeigt die durchschnittliche jährliche Veränderung der Arbeitsproduktivität für zwei Perioden, für die Jahre 2001 bis 2007 und für die Jahre 2010 bis 2017. Diese durchschnittlichen Veränderungsraten werden jeweils zerlegt in Intensivierung des ICT-Kapitals, Intensivierung des Nicht-ICT-Kapitals und Veränderung der Multifaktorproduktivität. Dabei ergeben hier die Summe aus Multifaktorproduktivität und Intensivierungen die Veränderungsrate der Arbeitsproduktivität (dargestellt als Punkt; als Summe dieser Faktoren).

[27] OECD (2019).

[28] Wiegmann (2008).

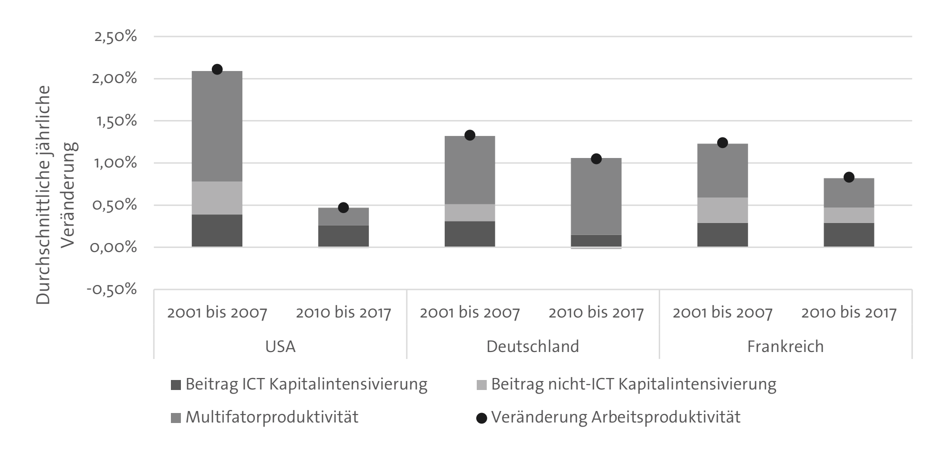


Abb. 5: Zerlegung der Arbeitsproduktivität in Kapitalintensivierung und Multifaktorproduktivität
Quelle: OECD (2019).

Für die Periode 2001 bis 2007 verzeichnen alle hier verglichenen Volkswirtschaften positive Impulse aus allen Zerlegungsfaktoren. In den USA und Frankreich fallen die Beiträge der Intensivierung des ICT-Kapitals und des übrigen Kapitals jeweils beinahe gleich aus. In Deutschland scheint das ICT-Kapital einen wichtigeren Beitrag zu leisten als das Nicht-ICT-Kapital. In allen Wirtschaften erklärt der Beitrag der Multifaktorproduktivität mindestens die Hälfte der Veränderung der Arbeitsproduktivität.

In der Periode 2010 bis 2017 ist das Bild weniger homogen. Die Multifaktorproduktivität bleibt in allen Volkswirtschaften wichtig, doch sie ist unterschiedlich wichtig. In Deutschland erklärt sie fast 90 % der Veränderung der Arbeitsproduktivität, in den anderen Ländern zwischen 30 und 45 %. Die Intensivierung des Nicht-ICT-Kapitals leistet in den USA und in Deutschland einen leicht negativen Beitrag; in Frankreich gehen hingegen 21 % des Wachstums der Arbeitsproduktivität auf die Intensivierung des Nicht-ICT-Kapitals zurück. In allen hier betrachteten Volkswirtschaften trägt die Intensivierung des ICT-Kapitals zur Erhöhung der Arbeitsproduktivität bei. In den USA ist dieser Beitrag sogar der größte in dieser Zerlegung, das heißt die Intensivierung des ICT-Kapitals leistet einen größeren Beitrag zur Erklärung der Arbeitsproduktivität als die Multifaktorproduktivität.

Gemäß diesen Zahlen geht also die Entwicklung der Arbeitsproduktivität in Deutschland fast vollständig auf die Multifaktorproduktivität zurück – das

heißt auf Faktoren wie die Qualität des Managements, Entscheidungen zum Einsatz der Arbeit sowie die Kombination mit dem Kapital oder die Prozessgestaltung. Wenn der Anstieg der Produktivität während des ersten Halbjahres 2020 hauptsächlich mit der Multifaktorproduktivität zu tun hatte, dann hat in Deutschland die Entwicklung der Arbeitsproduktivität schon seit langem und vor allem mit dieser Multifaktorproduktivität zu tun.

Schlussfolgerung

Was ist das Ergebnis nun? Generell lässt sich sagen:

- Die Messung der Arbeitsproduktivität, so wie sie heute Usus ist, zeigt eine nachlassende Dynamik.
- Während des Lockdowns im ersten Halbjahr 2020 konnten hingegen Produktivitätsverbesserungen umgesetzt werden.
- Diese gehen vermutlich mit der Ausnutzung von punktueller Flexibilisierung, Deregulierung, Digitalisierung und selbstgesteuerter Arbeit einher, also gehen sie auf eine höhere Multifaktorproduktivität zurück.
- Diese Erklärung wird auch dadurch plausibilisiert, dass in Deutschland die Multifaktorproduktivität beinahe 90 % der Arbeitsproduktivität erklärt.
- In Deutschland zeigt die Multifaktorproduktivität tendenziell eine Verbesserung über die Zeit an, was im Kontrast zu ihrer Entwicklung etwa in den USA und Frankreich steht.

Und trotzdem sind nicht alle Fragen beantwortet worden. Das Gegenteil ist der Fall. Die Betrachtung der Entwicklung der Arbeitsproduktivität unter Rückgriff auf die Multifaktorproduktivität führt zu weiteren Fragen, welche allesamt Forschungsdesiderate sind:

- Wie ist die Multifaktorproduktivität zu interpretieren? Wie kann sie gelesen werden, sodass sie die Verbesserungen der Managementqualität anzeigt?
- Welche Rolle spielen „Intangibles" für die Multifaktorproduktivität? Wenn sie gemäß der „Mismeasurement"-Hypothese von der Erfassung der Wertschöpfung ausgeschlossen werden, wie ist es dann möglich, dass das Wachstumsresiduum sie dennoch einschließt?
- Wenn die Multifaktorproduktivität streng rechnerisch eine Residualgröße ist, warum erklärt sie ausgerechnet in Deutschland beinahe 90 % der Arbeitsproduktivität?

– Welche anderen, präziseren Wege gibt es, die Multifaktorproduktivität zu interpretieren?
– Selbst dann, wenn die Multifaktorproduktivität nicht eine Lösung für die „Mismeasurement"-Hypothese ist: Wie kann sie nützlich gemacht werden, um das Problem besser zu verstehen?[29]

Literatur

Ahmed, T. und Bhatti, A. A. (2020): Measurement and Determinants of Multi-Factor Productivity: A Survey Of Literature, Journal of Economic Surveys 34(2), S. 293–319.

Baumol, W. J. und Bowen, W. G. (1966): Performing Arts: The Economic Dilemma: A Study of Problems Common to Theater, Opera, Music and Dance, The Twentieth Century Fund, New York.

Burda, M. C. (2018): Gesamtwirtschaftliche Arbeitsproduktivität, IZA World of Labor, April, S. 435–445.

Cantner, U., Krüger, J. und Hanusch, H. (2007): Produktivitäts- und Effizienzanalyse: Der nichtparametrische Ansatz, Springer-Verlag, Berlin et al.

Cardona, M., Kretschmer, T. und Strobel, T. (2013): ICT and productivity: conclusions from the empirical literature, Information Economics and Policy, 25(3), S. 109–125.

Corrado, C., Hulten, C. und Sichel, D. (2009): Intangible Capital and US Economic Growth, Review of Income and Wealth 55(3), S. 661–685.

Corrado, C. A. und Hulten, C. R. (2010): How Do You Measure a „Technological Revolution"?, American Economic Review 100(2), S. 99–104.

Erken, H., Donselaar, P. und Thurik, R. (2018): Total factor productivity and the role of entrepreneurship, The Journal of Technology Transfer, 43(6), S. 1493–1521.

Gordon, R. J. (2016): The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War, Princeton University Press, Princeton.

Jorgenson, D. W. (2017): World KLEMS: Productivity and Economic Growth in the World Economy: An Introduction, International Productivity Monitor (33), S. 1–8.

Karremans, J. (2020): Political alternatives under European economic governance: evidence from German budget speeches (2009–2019), Journal of European Public Policy 27(11), S. 1599–1621.

[29] Eine neuere Arbeit, welche in diese Richtung geht, ist Zimmermann (2020).

Kuntze, P. und Mai, C. M. (2020): Arbeitsproduktivität: Nachlassende Dynamik in Deutschland und Europa, WISTA–Wirtschaft und Statistik 72(2), S. 11–24.

Lemieux, P. (2018): What's Wrong with Protectionism: Answering Common Objections to Free Trade, Rowman & Littlefield, London.

Schneider, H. (2020): Das Rätsel der Produktivität, Springer-Gabler-Verlag, Wiesbaden.

Syverson, C. (2017): Challenges to Mismeasurement Explanations for the US Productivity Slowdown, Journal of Economic Perspectives 31(2), S. 165–86.

OECD (2015): The Future of Productivity, OECD Publishing, Paris.

OECD (2019): OECD Productivity Statistics 2019, OECD Publishing, Paris, DOI https://doi.org/10.1787/g2g99a92-en.

OECD (2020a): Labour Productivity and Utilisation (indicator), DOI https://doi.org/10.1787/02c02f63-en.

OECD (2020b): Multifactor Productivity (indicator), DOI https://doi.org/10.1787/a40c5025-en.

Van Ark, B., Hao, J. X., Corrado, C. und Hulten, C. (2009): Measuring Intangible Capital and Its Contribution to Economic Growth in Europe, EIB Papers 14(1), S. 62–93.

Van Ark, B. (2016): The Productivity Paradox of the New Digital Economy, International Productivity Monitor 31(Fall), S. 3–18.

Wiegmann, J. (2008): Produktivitätsentwicklung in Deutschland, Peter Lang, Bern.

Zimmermann, M. (2020): Immaterielles Kapital und Produktivität im Verarbeitenden Gewerbe: Ergebnisse auf Basis verknüpfter Einzeldaten (Micro Data Linking), WISTA-Wirtschaft und Statistik, 72(3), S. 61–75.

Teil 2: Neue empirische Verfahren für die Datenökonomie

Nowcast als Forecast

Neue Verfahren der BIP-Prognose in Echtzeit

Christina Heike Maaß

Einleitung

(Finanz-)politische Entscheidungen werden meist nach Bewertung aktueller und zukünftig erwarteter ökonomischer Entwicklungen getroffen.[1] Dafür benötigen Entscheidungsträger:innen aus der (Wirtschafts-) Politik beziehungsweise aus den Zentralbanken möglichst aktuelle Konjunkturdaten, um bestmöglich Einfluss auf die gegenwärtige Wirtschaftslage ausüben zu können.[2] Da wichtige volkswirtschaftliche Kennzahlen wie das Bruttoinlandsprodukt (BIP) zumeist nur in mehrmonatigen Intervallen und mit Verzögerung veröffentlicht werden, ist der Großteil der etablierten Indikatoren in Zeiten beschleunigter wirtschaftlicher Veränderungen nicht mehr agil genug.[3] Deswegen befassen sich Ökonom:innen tiefgehend mit der Verbesserung des makroökonomischen Monitorings in Echtzeit, um ein Verfahren zu entwickeln, mit dem die Gegenwart und die nahe Vergangenheit prognostiziert werden können.[4]

Eine Prognose des gegenwärtigen Zustands beziehungsweise der nahen Zukunft oder Vergangenheit, am Rande der verfügbaren Daten, wird als „Nowcast" bezeichnet.[5] Dieser Begriff setzt sich aus den englischen Wörtern *now* (jetzt) und *forecast* (Prognose) zusammen. Er bedeutet das Beobachten der aktuellen Wirtschaftslage in Echtzeit durch Prognose der Gegenwart, wobei die gegenwärtige

---

1 Giannone/Reichlin/Small (2008), S. 665.

2 BMWi (2019), S. 12; Giannone/Reichlin/Small (2008), S. 665; Cepni/Guney/Swanson (2020), S. 18.

3 Aastveit et al. (2014), S. 48; Giannone/Reichlin/Small (2008), S. 665; Hauf/Stehrenberg/Zwick (2020), S. 52.

4 Aastveit et al. (2014), S. 48; Stock/Watson (2017), S. 72.

5 Armstrong (2004), S. 798; Bańbura et al. (2013), S. 196; Sachverständigenrat (2019), S. 54.

Prognose immer wieder aktualisiert wird.[6] Abstrakter formuliert, besteht beispielsweise der Nowcast des BIP-Wachstums eines Landes zum Zeitpunkt *t*, dargestellt durch die Variable y_t^Q, aus der orthogonalen Projektion dieser Variablen auf das Informationsset Ω_v:

$$P\left[y_t^Q \mid \Omega_v\right] = E\left[y_t^Q \mid \Omega_v\right]$$

Das Informationsset enthält alle zum Zeitpunkt *v* verfügbaren Informationen, wobei *v* sich auf das Datum einer bestimmten Datenveröffentlichung bezieht. Dabei stellt $E[\cdot \mid \Omega_v]$ den bedingten Erwartungswert dar.[7]

Datenökonomie und künstliche Intelligenz (KI) eröffnen auch für Prognosen der Wirtschaftsleistung neue Möglichkeiten. Als Datenquelle für die Erstellung tagesaktueller Vorhersagen eignet sich vor allem der mit der neuen Datenökonomie einhergehende Bereich Big Data, da diese Daten für eine Vielzahl von Aktivitäten in Echtzeit und in großen Mengen zur Verfügung stehen. Für die bestmögliche Analyse von Big Data im Bereich Nowcasting erweist sich dabei insbesondere die Anwendung von KI in Form des maschinellen Lernens (ML) als effizient, da die Methoden des ML die Komplexität und Größe dieser Datenart am besten verarbeiten.[8]

Abbildung 1 vermittelt einen ersten Eindruck von der Genauigkeit von Nowcasts. Für den Zeitraum 1.1.2019 bis 31.12.2020 wird der BIP-Nowcast der Federal Reserve Bank of New York[9] für die USA geplottet (dunkelgrau). Demgegenüber werden die tatsächlichen Wachstumsraten des US-amerikanischen BIP als hellgraue Punkte zum jeweiligen Ende des Quartals[10] dargestellt. Diese Werte stammen vom *Bureau of Economic Analysis* (BEA)[11] der US-Regierung. Es wird ersichtlich, dass der Nowcast die BIP-Entwicklung auch im Verlauf der Coronapandemie (seit Mitte März 2020) relativ gut abbildet. Die Abweichungen in dieser unsicheren Periode sind allerdings um einiges deutlicher als vorher. Eine genauere Beschreibung der Grafik erfolgt im weiteren Verlauf dieses Beitrags.

---

[6] Federal Reserve Bank of New York (2019), S. 66.

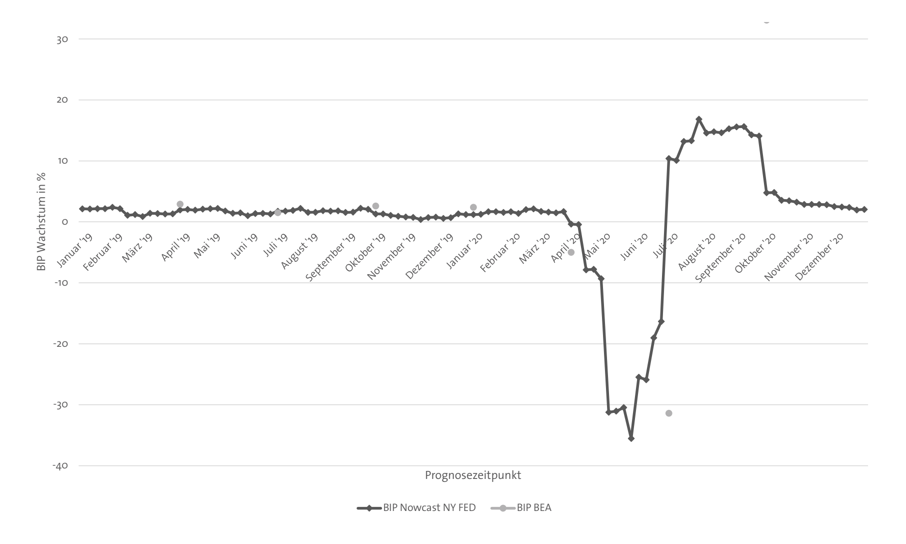
[7] Bańbura/Giannone/Reichlin (2011), S. 196.

[8] Varian (2014), S. 6.

[9] Federal Reserve Bank of New York (2020).

[10] Zum Verständnis dessen auf welches Quartal sich das tatsächliche BIP des Bureau of Economic Analysis (BEA) bezieht, wird das tatsächliche BIP am Ende des jeweiligen Quartals anstelle am Tag der tatsächlichen Veröffentlichung (einige Wochen später) dargestellt.

[11] BEA (2020).



Anmerkung: NY FED = Federal Reserve Bank of New York, BEA = Bureau of Economic Analysis.

Abb. 1: US-BIP-Wachstum: NY-FED-Nowcast versus BEA 01/19–12/20
Quelle: Eigene Darstellung, basierend auf BEA (2020) und Federal Reserve Bank of New York (2020).

Dieser Beitrag bietet einen Überblick über die gegenwärtige Nowcast-Literatur und untersucht die Möglichkeiten und Grenzen der Erstellung (tages-) aktueller Prognosen (sogenannter Nowcasts) der gegenwärtigen Wirtschaftslage mit Big Data und ML. Zum Vergleich der verschiedenen Prognosequalitäten wird das statistisch-empirische Kriterium Wurzel der mittleren Fehlerquadratsumme (*root mean square error* [RMSE]) angelegt.[12]

Im folgenden Abschnitt werden zunächst die Eigenschaften des Nowcasts genauer beschrieben und anschließend „herkömmliche“ Anwendungsmöglichkeiten dieser Methode erläutert. Der nächste Abschnitt geht darauf ein, was unter dem Begriff Big Data verstanden wird sowie welche Datenquellen aus dem Bereich Big Data stammen, und umreißt die Herangehensweise von Verfahren des maschinellen Lernens. Anschließend werden Möglichkeiten und Grenzen des Nowcasts mit Big Data und maschinellem Lernen diskutiert. Schließlich werden die Ergebnisse dieses Beitrags zusammengefasst und es wird ein Ausblick auf potenzielle zukünftige Entwicklungen gegeben.

[12] Pons (2000), S. 54.

Tagesaktuelle Prognosen – Nowcasts

Nowcasts sind Prognosen der Gegenwart. Sie unterscheiden sich von anderen Prognoseverfahren jedoch nicht nur dadurch, dass sie anstelle der Zukunft die gegenwärtige Lage prognostizieren, sondern auch durch die Asynchronität der verwendeten Datenreihen. Die verschiedenen Indikatoren, auf Grundlage derer ein Nowcast erstellt wird, werden zu unterschiedlichen Zeitpunkten und mit unterschiedlichen Frequenzen veröffentlicht – manche Variablen werden monatlich veröffentlicht, andere wöchentlich oder sogar täglich –, sodass der zugrundeliegende Datensatz unausgewogen ist.[13] Da für einen Nowcast immer die aktuellsten Kennzahlenwerte genutzt werden, ist der Zeitpunkt der letzten verfügbaren Beobachtung in jeder Datenserie verschieden. Daher hat der Gesamtdatensatz, das Informationsset Ω, eine spezielle Struktur, die sich „gezackte Kante" (*ragged* beziehungsweise *jagged edge*) nennt.

Des Weiteren wird für die Erstellung eines Nowcasts nicht nur eine Prognose angefertigt, sondern eine ganze Reihe von Prognosen. Die erste Vorhersage wird mit nur ganz wenigen Informationen angefertigt. Sobald neue Daten vorliegen, wird die vorherige Schätzung aktualisiert und ein neuer, präziserer Nowcast erstellt. Es wird also eine Sequenz von Projektionen $E[y_t^Q|\Omega_v], E[y_t^Q|\Omega_{v+1}]$ angefertigt, bei der $v, v+1$ sich auf die Zeitpunkte aufeinanderfolgender Datenveröffentlichungen beziehen. Normalerweise liegen diese Zeitpunkte dicht beieinander und ändern sich. Deswegen hat v eine hohe Frequenz und ist unregelmäßig verteilt. Die neueren Prognosen enthalten dabei sowohl mehr Informationen als auch Revisionen vorheriger Daten. Der bestehende Nowcast ändert sich nur, wenn die neue Prognose von der alten abweicht, da unerwartete Informationen, sogenannte *News*, veröffentlicht wurden.[14] Allerdings geht Nowcasting weit über einzelne Prognosen für einen bestimmten Zeitraum hinaus. Ziel ist es viel mehr, einen Rahmen zu entwickeln, der in Echtzeit den Fluss an Datenveröffentlichungen lesen kann.[15] Deswegen ist Nowcasting nicht nur gleichzeitiges Forecasting, da die Nutzung und das Timing gleichzeitiger Daten bedeutend andere Komponenten mit sich bringen als vorab Forecasting.[16]

---

[13] Giannone/Reichlin/Small (2008), S. 666; Bańbura/Giannone/Reichlin (2011), S. 196.

[14] Bańbura/Giannone/Reichlin (2011), S. 196–197.

[15] Bańbura et al. (2013), S. 207.

[16] Castle/Fawcett/Hendry (2009), S. 87.

Nowcasts können für sämtliche niederfrequenten Variablen angefertigt werden, die mit einiger Verzögerung veröffentlicht werden, für die jedoch höherfrequente, frühzeitigere Daten verfügbar sind.[17] Ursprünglich wurden Nowcasts hauptsächlich in der Meteorologie zur minutengenauen Prognose von Gewittern verwendet.[18] In den letzten Jahren ist die Relevanz dieses Verfahrens jedoch auch in der Ökonomie deutlich gestiegen. Dies zeigt sich beispielsweise an den verstärkten Aktivitäten statistischer Ämter, Institute und Zentralbanken.[19]

Man benötigt Nowcasts in den Wirtschaftswissenschaften, da wichtige volkswirtschaftliche Kennzahlen wie das BIP nur vierteljährlich berechnet werden und erste offizielle BIP-Schätzungen des vergangenen Quartals in vielen Ländern erst vier (USA) bis sechs (Eurozone) Wochen nach Ende des Quartals veröffentlicht werden. Somit muss in der Eurozone bis circa Mitte Mai (beziehungsweise Ende April für eine Schnellschätzung) gewartet werden, um erste Prognosen für das aktuelle Jahr zu erhalten.[20] Aufgrund des immensen volkswirtschaftlichen Interesses am Wirtschaftswachstum liegt der Schwerpunkt dieses Kapitels auf der Prognose des BIP.

Um das gegenwärtige BIP zu prognostizieren, werden früher beziehungsweise häufiger als das BIP veröffentlichte Variablen verwendet, im Idealfall sogar Echtzeitdaten.[21] Zum Beispiel wird die Industrieproduktion der Eurozone im Januar schon Mitte März veröffentlicht. Dieser Indikator misst einen wichtigen Bestandteil des BIP und ist von großer Bedeutung für dessen kurzfristige Entwicklung.[22] Für Deutschland sind außerdem beispielsweise Auftragseingänge im Verarbeitenden Gewerbe, Produktion im Produzierenden Gewerbe oder der ifo Geschäftsklimaindex wichtige Frühindikatoren.[23] Diese Kennzahlen zählen zu den „harten" Faktoren. Noch kurzfristigere, möglicherweise aber auch ungenauere Informationen kann man „weichen" Faktoren wie aktuellen Umfragewerten entnehmen. Werte für BIP-Erwartungen werden häufig schon vor Ende des betreffenden Monats veröffentlicht. Zusätzlich können auch sämtliche an-

---

[17] Bańbura/Giannone/Reichlin (2011), S. 194.

[18] siehe Deutscher Wetterdienst (2020) und World Meteorological Organization (2017).

[19] Siehe Deutsche Bundesbank (2020); Federal Reserve Bank of New York (2020); ifo Institut (2021); Statistisches Bundesamt (2020).

[20] Bańbura/Giannone/Reichlin (2011), S. 193–194; Bańbura et al. (2013), S. 196.

[21] Bańbura et al. (2013), S. 196; Cepni/Guney/Swanson (2020), S. 18.

[22] Bańbura/Giannone/Reichlin (2011), S. 193.

[23] BMWi (2019), S. 12.

deren Datenveröffentlichungen von Bedeutung sein und die Märkte beeinflussen,[24] beispielsweise Finanzmarktdaten wie Zinsen, Aktien- oder Wechselkurse sowie nachfrageseitige Variablen wie der Stromverbrauch. Darüber hinaus eignet sich für die Erstellung tagesaktueller Nowcasts insbesondere die Verarbeitung großer Informationsmengen im Bereich Big Data wie Social-Media-Aktivitäten oder Mobilfunkdaten. Diese Daten können theoretisch in Echtzeit in aktuelle Prognosen einfließen. Möglichkeiten und Grenzen dieses Ansatzes werden in einem nachfolgenden Abschnitt dieses Beitrags diskutiert.

Technisch gesehen wurden beim Nowcast des BIP-Wachstums im gegenwärtigen Quartal ursprünglich häufig einfache, kleinmaßstäbige Modelle verbunden, sogenannte Brückengleichungen (*bridge equations*). Dabei werden die kleinen Modelle genutzt, um die Informationen, die in einer oder mehreren monatlichen erklärenden Schlüsselvariablen (zum Beispiel Werte für die Industrieproduktion oder Umfrageergebnisse) enthalten sind, mit der vierteljährlichen Wachstumsrate des BIP zu koppeln, die nach den monatlichen Daten veröffentlicht wird.[25] Anhand eines Gleichungssystems wird eine Verbindung zwischen der zu prognostizierenden Variablen und den erklärenden Variablen hergestellt. Dabei wird überbrückt, dass die Daten alle zu unterschiedlichen Zeitpunkten veröffentlicht werden.[26]

Allerdings werden für Nowcasts idealerweise komplexere Modelle verwendet, bei denen auch die Kommentierung und Interpretation der überarbeiteten Schätzungen berücksichtigt wird. Mit den komplexeren Modellen wird die Veränderung des BIP nicht nur quantitativ prognostiziert, sondern es wird auch qualitativ bestimmt, welche Faktoren hauptsächlich verantwortlich sind. Dafür muss die gemeinsame Dynamik der monatlichen Inputvariablen und der vierteljährlichen Zielvariablen in einem einheitlichen Rahmen modelliert werden.[27] Prinzipiell kann dafür jedes dynamische Modell verwendet werden, das mit verschiedenen Frequenzen und fehlenden Daten umgehen kann und die gemeinsame Dynamik der erklärenden und der Zielvariablen abbilden kann.[28] In der Praxis werden häufig Faktormodelle oder vektorautoregressive Modelle (VAR-Modelle) verwendet.[29]

---

[24] Bańbura/Giannone/Reichlin (2011), S. 193–194.

[25] Giannone/Reichlin/Small (2008), S. 666.

[26] Deutsche Bundesbank (2018), S. 19–20.

[27] Bańbura/Giannone/Reichlin (2011), S. 194.

[28] Bańbura/Giannone/Reichlin (2011), S. 212.

[29] Sachverständigenrat (2019), S. 54; Deutsche Bundesbank (2018), S. 15.

Als statistisch-empirische Kriterien für den Vergleich und die Bewertung der Prognosegüte verschiedener Verfahren werden traditionell der mittlere absolute Fehler (*mean absolute error* = MAE), die Wurzel der mittleren Fehlerquadratsumme (*root mean square error* [RMSE]) und Theils U angelegt.[30] Da in den für diesen Beitrag relevanten Papieren überwiegend der RMSE angegeben wird, wird nachfolgend die Güte der verschiedenen Nowcasts anhand dieses Kriteriums bewertet. Der RMSE[31] wird durch die Abweichung der prognostizierten Werte von der Realität berechnet. Diese Abweichung wird quadriert, um negative und positive Differenzen gleich zu bewerten, es wird der Mittelwert gebildet und die Quadratwurzel gezogen. Im Verlauf des Beitrags wird daher immer wieder auf den RMSE eingegangen, um einen Überblick über die Prognosegüte der verschiedenen Ansätze zu erhalten. Am Ende des Beitrags werden alle RMSE in einer Tabelle einander gegenübergestellt.

Anwendungsmöglichkeiten von Nowcasts

Nowcasts sind für eine Vielzahl volkswirtschaftlicher Themengebiete relevant, wie Gesundheit, Produktion beziehungsweise Produktivität, Konsumentenverhalten und Wirtschaftswachstum. Bei der Formalisierung der Problemstellungen wird im traditionellen Sinne sowohl auf Partialmodelle wie Brückengleichungen und MIDAS-Typ-Gleichungen als auch auf gemeinsame Modelle wie Faktormodelle und Gemischte-Frequenz-VAR-Modelle zurückgegriffen. Eine Besonderheit der VAR-Modelle ist, dass die verwendeten Variablen von ihren vergangenen Werten sowie den vergangenen Werten der anderen Variablen abhängen.[32]

In einem frühen Standardwerk zur Problematik der zeitverzögerten Berichterstattung[33] untersuchte Jerald Lawless im Jahr 1994 die Berichterstattung bei meldepflichtigen Krankheiten. Anhand eines stationären *Maximum-Likelihood*-Verfahrens mit multinominalem Modell sowie eines nichtstationären *Random-Effects*-Modells konnte er Trends in der verzögerten Berichterstattung iden-

---

[30] Pons (2000), S. 54.

[31] Berechnung: RMSE = $\sqrt{\frac{1}{n} * \sum_{i=1}^{n} (y_i - f_i)^2}$, y_i = tatsächliche Werte, f_i = prognostizierte Werte (Gupta/Minai 2019), S. 372.

[32] Deutsche Bundesbank (2018), S. 18.

[33] Englisch: „*occured but not reported events*“.

tifizieren und somit bessere Prognosen bezüglich des Verlaufs von Krankheiten erstellen.[34] Dies ist insbesondere bei plötzlichen Ausbrüchen essentiell.[35]

Deswegen spielt das Nowcasting tatsächlicher Infiziertenzahlen auch in der aktuellen COVID-19-Pandemie eine große Rolle. Auch die Registrierung neuer SARS-CoV-2-Infektionen erfolgt oftmals zeitverzögert, da einerseits Infektions- und Meldedatum nicht übereinstimmen und andererseits das tatsächliche Infektionsdatum nicht bekannt beziehungsweise nicht zu bestimmen ist. Um trotzdem möglichst genaue Prognosen des aktuellen Verlaufs der Krankheit treffen zu können, können die fehlenden Informationen bezüglich des tatsächlichen Infektionsdatums mit einem Nowcast geschätzt werden.[36] Nur wenn eine genaue Prognose der aktuellen Infektionszahlen vorliegt, kann ein eindeutiger Zusammenhang zwischen Ursache, zum Beispiel einer Veränderung der Coronaregeln, und Wirkung, beispielsweise höheren oder niedrigeren Infektionszahlen, hergestellt werden. Das Robert Koch-Institut (RKI) schätzt daher fehlende Datenpunkte basierend auf den statistischen Zusammenhängen der verfügbaren Daten (sogenannte multiple Imputation), um den Zeitpunkt der Erkrankung an SARS-CoV-2 zu schätzen und das aktuelle Infektionsgeschehen besser beurteilen zu können.[37]

Abgesehen von medizinischen Studien sind Nowcasts der gegenwärtigen Wirtschaftsleistung von großer Bedeutung und Gegenstand verschiedener Forschungsarbeiten. Aktuelle Veröffentlichungen untersuchen überwiegend Nowcasts, die in Pseudo-Echtzeit erstellt wurden. Es wird also simuliert, was ein Nowcast im jeweiligen Zeitraum prognostiziert hätte. Dafür werden nur Daten verwendet, die zum jeweiligen Prognosezeitpunkt verfügbar waren. Die erhaltenen Prognosen können dann direkt mit den tatsächlichen Werten verglichen werden. Dabei werden verschiedene Herangehensweisen gewählt.

In einer der neuesten Studien ermitteln Cepni/Guney/Swanson BIP-Nowcasts basierend auf Datensätzen ökonomischer Indikatoren des Informationsdienstleisters Bloombeg. Für fünf Schwellenländer (Brasilien, Indonesien, Mexiko, Südafrika, Türkei) werden dynamische Faktormodelle durch makroökonomische Unsicherheits- und Überraschungsfaktoren erweitert, die die intensiven globalen

[34] Lawless (1994).

[35] Höhle / An der Heiden (2014).

[36] Günther et al. (2020).

[37] An der Heiden / Hamouda (2020).

Verflechtungen der Volkswirtschaften abbilden.[38] Die Überraschungsfaktoren bestehen aus Abweichungen der Erwartungen professioneller Prognostiker von der Realität, die Unsicherheitsfaktoren enthalten Werte aus der Wirtschafts-, Handels- und Geldpolitik sowie Migration.[39] Cepni/Guney/Swanson kommen zu dem Ergebnis, dass unterschwellige Unsicherheits- und Überraschungsfaktoren einen signifikanten vorhersagenden Inhalt für die Wachstumsprognose von Schwellenländern haben.[40] Somit kann die Messung von Unsicherheit für eine genaue Vorhersage nützlich sein. Der RMSE ihres autoregressiven Standardmodells liegt je nach Land und den drei verwendeten Nowcast-Zeitpunkten zwischen 0,85 und 2,34. Außerdem zeigen sie, dass der Prognosegehalt bei der Konstruktion der Schrumpfungstypfaktoren mit LASSO, adaptivem LASSO, bayesianischem LASSO und adaptivem bayesianischem LASSO verbessert ist (niedrigerer RMSE).[41]

Im Gegensatz dazu berücksichtigen Aastveit et al. Prognoseunsicherheiten als Wahrscheinlichkeitsverteilung in sogenannten Dichte-Nowcasts des US-amerikanischen BIP.[42] Diese Nowcasts bestehen aus einem System von Brückengleichungsmodellen, Faktormodellen und vektorautoregressiven Modellen. Dafür verwenden Aastveit et al. mehr als 100 monatliche Indikatoren wie Finanzmarktdaten, Verbraucherpreise und Umfragewerte von der Federal Reserve Bank in St. Louis. Die verschiedenen Dichte-Nowcasts werden zunächst zu drei Nowcasts je Modellklasse und schließlich zu einem Gesamt-Nowcast kombiniert. Die Prognosen werden mit jeder Datenveröffentlichung besser und der kombinierte Nowcast liefert gemäß dem von den Autoren verwendeten *log score*[43] die besten Prognosen.[44]

Nowcasts des BIP in der Eurozone im Jahr 2008 finden sich in Bańbura/Giannone/Reichlin.[45] Datengrundlage sind 26 „wichtige“ Indikatoren aus der Eurozone mit höchstens monatlicher Frequenz. Die Autoren vergleichen die Nowcasts eines dynamischen Faktormodells mit denen eines univariaten nai-

---

[38] Cepni/Guney/Swanson (2020), S. 33.

[39] Cepni/Guney/Swanson (2020), S. 19.

[40] Cepni/Guney/Swanson (2020), S. 32.

[41] Cepni/Guney/Swanson (2020).

[42] Aastveit et al. (2014).

[43] Der *log score* ist der Logarithmus einer Wahrscheinlichkeitsdichtefunktion, die an der Stelle des prognostizierten Ergebnisses ausgewertet wird und somit eine Maßzahl für die Passform der Dichtefunktion angibt. Aastveit et al. (2014), S. 51.

[44] Aastveit et al. (2014), 48 und 59.

[45] Bańbura/Giannone/Reichlin (2011).

ven Random-Walk-Modells. Während das dynamische Faktormodell zu Beginn der Prognoseperiode einen RMSE von 0,28 aufweist, der im Laufe des Quartals auf 0,18 sinkt, liegt der RMSE des Random-Walk-Modells vergleichsweise konstant bei etwas über 0,30. Des Weiteren ergibt ihre Analyse, dass weiche Faktoren wie Umfrageergebnisse zu Beginn eines Quartals hohen Erklärungsgehalt haben. Im weiteren Verlauf des Quartals, mit der Veröffentlichung harter Faktoren, sinkt ihre Bedeutung jedoch.[46]

Bańbura/Giannone/Modugno/Reichlin erstellen Nowcasts für das US-amerikanische BIP in Echtzeit.[47] Dabei verwenden sie an dem Tag, an dem sie eine Prognose erstellen, alle bis dahin vorliegenden Daten der interessierenden Variablen. Der Echtzeitdatensatz basiert auf Werten des Informationsdienstleisters Bloomberg und wird anhand von Daten der Federal Reserve Bank in St. Louis evaluiert. In diesem Paper werden sowohl gemeinsame Modelle – ein Faktormodell mit täglichen Daten, ein analoges Modell mit monatlicher Frequenz sowie ein kleines Faktormodell mit fünf „harten" Indikatoren[48] – als auch Brückengleichungen als Partialmodell verwendet. Die Autoren finden heraus, dass tagesaktuelle Finanzdaten nur einen geringen Erklärungsgehalt haben und stattdessen monatliche Veröffentlichungen der Werte makroökonomischer Variablen entscheidend sind.[49] Das Brückengleichungsmodell hat den höchsten RMSE zwischen 0,58 und 0,66, gefolgt von dem Modell mit den fünf harten Faktoren (RMSE zwischen 0,45 und 0,63). Eine noch höhere Prognosegüte weist das Modell mit täglicher Frequenz auf (RMSE zwischen 0,45 und 0,60) und den niedrigsten RMSE zwischen 0,43 und 0,57 erzielt das Faktormodell mit monatlichen Daten.[50]

Abgesehen von diesen akademischen Publikationen beschäftigen sich diverse Zentralbanken und (Forschungs-)Institute mit der Entwicklung eigener Nowcasts. Dabei bedienen sie sich sowohl herkömmlicher ökonomischer Indikatoren als auch hochfrequenter Big-Data-Datensätze. Beispielsweise veröffentlicht die Federal Reserve Bank of New York ihren wöchentlich aktualisierten

---

[46] Bańbura/Giannone/Reichlin (2011), S. 198, 205, 209 und 210.

[47] Bańbura et al. (2013).

[48] Reales BIP, Industrieproduktion, verfügbares Realeinkommen, Einzelhandelsumsatz und Beschäftigung.

[49] Bańbura et al. (2013), S. 214 sowie 218-219.

[50] Bańbura et al. (2013), S. 217.

New York Fed Staff Nowcast[51] zur Entwicklung des BIP in den USA.[52] Bei dieser Prognose verwendet sie ein dynamisches Faktormodell.[53] Abbildung 1 zeigt zwischen Januar 2019 und Dezember 2020 den Verlauf der Nowcasts gegenüber den tatsächlichen BIP-Werten, die ein paar Wochen nach Ablauf des jeweiligen Quartals veröffentlicht werden. Bereits vor Ende des ersten Quartals 2020 weist der Nowcast eine Abwärtstendenz auf, die sich nach Veröffentlichung der tatsächlichen BIP-Entwicklung und dem Fortschreiten der COVID-19-Pandemie deutlich verstärkt. Das zweite Quartal 2020 wird durch den Nowcast ziemlich gut vorhergesagt, insbesondere in Anbetracht der sehr unsicheren Lage. Die Erholung im dritten Quartal 2020 wurde deutlich niedriger prognostiziert als sie tatsächlich ausgefallen ist. Die Prognose für das vierte Quartal ist im positiven, einstelligen Bereich. Die New Yorker Zentralbank erwartet für das Ende des Jahres ein BIP-Wachstum zwischen 2 und 5 %. Das tatsächliche BIP-Wachstum gemäß BEA wurde erst nach dem Druck dieses Artikels veröffentlicht. Es bleibt festzuhalten, dass der Nowcast der New Yorker Zentralbank auch im sehr unruhigen Jahr 2020 gute Ergebnisse geliefert hat.

Auch die Federal Reserve Bank of Atlanta prognostiziert mit ihrem GDPNow Nowcast anhand von Brückengleichungen das US-amerikanische BIP. Dabei werden die ersten Schätzungen des aktuellen Quartals ungefähr einen Monat nach Ablauf des vorherigen Quartals erstellt.[54] Die *out of sample* Prognosegüte des GDPNow-Modells einen Tag vor der Veröffentlichung der ersten BIP-Schätzung liegt bei einem RMSE von 1,15. 85 Tage vor der Veröffentlichung liegt er noch bei durchschnittlich 2,2.[55]

Für die EU und die Eurozone veröffentlicht Eurostat seit 2016 BIP-Schnellschätzungen 30 Tage ($t+30$) nach Quartalsende.[56] Diese Prognose besteht aus den jeweiligen $t+30$ Prognosen verschiedener Mitgliedsländer sowie von Eurostat angefertigten Schätzungen des BIP-Wachstums der Länder mit fehlenden Daten. Für diese werden mit Hilfe der Daten für die ersten beiden Monate des Quartals mehrere Nowcasts angefertigt. Darin werden normalerweise die Wachstumsraten des Industrieproduktionsindex, des deflationierten Einzel-

---

[51] Abrufbar unter https://www.newyorkfed.org/research/policy/nowcast.html.

[52] Bok et al. (2018), S. 616.

[53] Federal Reserve Bank of New York (2019), S. 69.

[54] Federal Reserve Bank of Atlanta (2020).

[55] Higgins (2014, 25 und 28).

[56] Eurostat (2019).

handels und des Umsatzes von Dienstleistungen anhand der Gewichtungen der volkswirtschaftlichen Gesamtrechnung des jeweiligen Landes verarbeitet. Anhand dieser Indikatoren sowie von vergangenen BIP-Werten und des Indikators für die wirtschaftliche Stimmung werden dann autoregressive Modelle mit verteilter Verzögerung (*autoregressive distributed lag* [ADL]) oder Modelle autoregressiver integrierter gleitender Mittelwerte mit exogenen Variablen (*autoregressive integrated moving average with exogenous variables* [ARIMAX]) geschätzt. Dieses Ergebnis geht mit in die Prognose für die gesamte EU beziehungsweise Eurozone ein.[57] Der RMSE für die Quartalsschätzung für den Zeitraum 2012 bis 2015 liegt in der EU bei 0,06 bis 0,08. In der Eurozone ist er minimal höher (0,07 bis 0,10).[58] Dieser Wert ist allerdings nur schwer mit den anderen vergleichbar, da es sich um eine Aggregation verschiedener Modellprognosen der einzelnen Länder handelt und somit kein eindeutiges Modell zugrunde liegt.

Viele große europäische Länder wie Frankreich, Italien, Spanien, Belgien und Österreich veröffentlichen diese frühe Eurostat-Schätzung ihres BIP bereits.[59] Für Deutschland publiziert das statistische Bundesamt seit dem zweiten Quartal 2020 eine BIP-Flash-Schnellschätzung, die ebenfalls 30 Tage (t + 30) nach Ablauf des Quartals zur Verfügung steht. Außerdem arbeitet es intern an einem BIP-Nowcast, der bereits zehn Tage ($t + 10$) nach Quartalsende erste Prognosen liefern könnte.[60] Das betreffende Verfahren beruht auf Modellen autoregressiver integrierter gleitender Mittelwerte (ARIMA).[61] Während der BIP-Flash $t + 30$ qualitativ sehr hochwertige Ergebnisse liefert, ist der BIP-Nowcast $t + 10$ zwar vielversprechend, aber noch deutlich anfälliger für Ungenauigkeit.[62]

Im November 2020 gab auch das ifo Institut bekannt, in Zukunft seine BIP-Schnellschätzung ifoCAST regelmäßig zusammen mit der Konjunkturprognose zu veröffentlichen. Für diesen Nowcast wurden zunächst aus über 300 Indikatoren mittels ML die passendsten ausgewählt. Anhand dieser wird zweimal monatlich ein dynamisches Faktormodell mit bayesianischen Methoden für das aktuelle Quartal geschätzt. Das neue Quartal wird ab Veröffentlichung der detaillierten Ergebnisse der Volkswirtschaftlichen Gesamtrechnung des

[57] Doody/Kokkinen/Wouters (2016), S. 13–14 und 16–17).

[58] Doody/Kokkinen/Wouters (2016), S. 25–26.

[59] Dickopf/Janz/Mucha (2019), S. 48.

[60] Dickopf/Janz/Mucha (2019), S. 48; Statistisches Bundesamt (2020).

[61] Dickopf/Janz/Mucha (2019), S. 50.

[62] Dickopf/Janz/Mucha (2019), S. 56.

Vorquartals geschätzt – also ab 55 Tagen nach Quartalsende. Die Prognosegüte des ifoCAST ist mit einem RMSE von 0,32 recht gut.[63]

Abgesehen davon entwickelt die Deutsche Bundesbank einen wöchentlichen Aktivitätsindex, mit dem sich die realwirtschaftliche Aktivität in Deutschland mit geringstmöglicher Zeitverzögerung bestimmen lässt. Er wird anhand eines *Expectation-Maximization*-Algorithmus berechnet und setzt sich aus hochfrequenten Indikatoren mit hohem Erklärungsgehalt aus den verschiedenen Wirtschaftsbereichen sowie dem BIP des vergangenen Quartals und der monatlichen Industrieproduktion zusammen.[64] Aktuelle Werte sind bei der Deutschen Bundesbank online abrufbar.[65]

Insgesamt werden für Nowcasts viele verschiedene Indikatoren verwendet. Zumeist liegt den Prognosen eine ganze Reihe von Kennzahlen bezüglich Leistungsbilanz, Beschäftigung, Finanzmärkten, Produktion und Konsum sowie unterschiedliche Umfragen zugrunde. Ein tiefergehender Vergleich der jeweils genutzten Kennzahlen würde den Umfang dieser Arbeit überschreiten und ist für einen ersten Überblick über verschiedene Nowcasting-Verfahren nicht entscheidend.

Big Data und Maschinelles Lernen

Auf Big Data basierende Datensätze wecken seit ein paar Jahren auch verstärkt das Interesse von Ökonom:innen. Verschiedene Studien zeigen, dass mit Big Data erstellte Prognosen besser als die herkömmlichen Prognosen sind, beispielsweise bei der Vorhersage des privaten Konsums[66] oder des BIP-Wachstums.[67] Dabei unterscheidet sich Big Data von herkömmlichen Daten durch enormes Volumen, hohe Geschwindigkeit, große Vielfalt, geringe Richtigkeit (im Sinne von Unsicherheiten in den Daten) und hohen Wert der Daten.[68] In anderen Worten stellt Big Data jenes Informationsgut dar, das durch sein großes Volumen sowie

---

[63] Lehmann/Reif/Wollmershäuser (2020), S. 31–32 und 38.

[64] Deutsche Bundesbank (2020).

[65] Abrufbar unter https://www.bundesbank.de/de/statistiken/konjunktur-und-preise/woechentlicher-aktivitaetsindex.

[66] Siehe Vosen/Schmidt (2011b).

[67] Siehe Woloszko (2017).

[68] Jin et al. (2015), S. 59.

seine große Geschwindigkeit und Vielfalt spezielle Technologien und analytische Verfahren erfordert, um es in (ökonomischen) Wert zu verwandeln.[69, 70] Dies liegt darin begründet, dass diese Daten unstrukturierter sind und höhere Dimensionen aufweisen, sodass das Ordnen der Daten und das Reduzieren der Dimensionen sehr komplex ist.[71] Daher ist die Nutzung von Big Data in vielerlei Hinsicht nicht mit der Nutzung herkömmlicher Daten vergleichbar.

Moderne Big-Data-Datensätze sind häufig in Echtzeit verfügbar und es gibt Daten zu vielen neuen Bereichen.[72] Ihre wichtigste Quelle sind im Internet erzeugte Daten. Dabei sind insbesondere Social-Media-Daten von großer Bedeutung.[73] Darüber hinaus kann Big Data auch aus Finanzmarktdaten, E-Commerce- und Scannerdaten, Mobilfunkdaten, Geolokalisierungsdaten, Such-(maschinen-)daten, Text- und Mediendaten, sensorischen Daten oder Transportdaten bestehen.[74] Praktisch jede elektronisch erfasste Handlung kann als Datengrundlage verwendet werden, was eine Fülle von Möglichkeiten für tagesaktuelle Prognosen eröffnet.

Aufgrund der besonderen Eigenschaften aktueller (Big-Data-)Datenströme, wie deren Größe und Komplexität, sollten für die Analyse von Big Data andere Methoden als die klassischen statistischen Verfahren verwendet werden.[75] Da herkömmliche Verfahren Daten in ein Modell einpassen, das den Zusammenhang zwischen Dateneingabe und -ausgabe möglichst gut beschreibt, kann das jeweilige Modell nicht besser sein als der angenommene Zusammenhang zwischen den Variablen selbst.[76] Die großen Datensätze ermöglichen jedoch flexiblere Beziehungen als einfache lineare Modelle. Diese können durch Techniken des maschinellen Lernens effektiv modelliert werden.[77] Bei diesen Verfahren muss keine Annahme über die funktionelle Beziehung der Variablen getroffen werden, sondern ein Algorithmus bestimmt die Funktion, die den

---

[69] „Big Data represents the Information assets characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value." Mauro/Greco/Grimaldi (2015), S. 103.

[70] Mauro/Greco/Grimaldi (2015), S. 103.

[71] Einav/Levin (2014b), S. 1.

[72] Einav/Levin (2014b), S. 1.

[73] Campos Vázquez/López-Araiza B. (2020), 129; Einav/Levin (2014b), S. 1.

[74] Garboden (2020), S. 8–18; Einav/Levin (2014b), S. 1.

[75] Varian (2014), S. 4 und 6.

[76] Jung/Patnam/Ter-Martirosyan (2018), S. 4.

[77] Varian (2014), S. 3.

Zusammenhang zwischen den Variablen am besten darstellt.[78] Anstelle einer linearen oder logistischen Regression erweisen sich nicht-lineare Verfahren wie Klassifikations- und Regressionsbäume, *Random Forests* oder bestrafte Regressionen (*penalized regressions*) wie LASSO, LARS und elastische Netze als besser geeignet.[79]

Das Vorgehen beim ML besteht aus zwei Elementen: (1) einer Lernmethode, bei der anhand von Daten die bestmögliche Einpassung der Inputvariablen bestimmt wird, und (2) einem Algorithmus, der die Beziehung zwischen Input und Output modelliert[80]. Eine Besonderheit des ML ist dabei der Fokus auf die Daten anstelle von Theorie.[81] Bei diesem Verfahren kann also mit einem *measurement without theory* nur auf Grundlage von Daten die Veränderung makroökonomischer Kennzahlen prognostiziert werden. Anschließend kann ML dann im Prinzip auch auf die zugrundeliegende Theorie schließen. Dies geschieht indem basierend auf den Ergebnissen der Algorithmen die richtige beziehungsweise best passende Theorie gesucht wird. Da mit Systemen des ML mehr und verschiedenartigere Informationen verarbeitet werden können, kann die Realität besser abgebildet und können somit potenziell bessere Prognosen erstellt werden.[82]

Allerdings kann bei der Anwendung mit Big Data die Vielzahl an vorhandenen Informationen zu einem sogenannten *overfitting* führen. Dabei werden Modelle durch unnötig viele erklärende Variablen zu komplex. Das Ergebnis ist, dass das Modell mit dem vorliegenden Datensatz sehr genaue Ergebnisse liefert, bei *out of sample*-Schätzungen allerdings schlecht abschneidet. Dies kann durch Bestrafung übermäßig komplexer Modelle, eine Unterteilung des Datensatzes in Training, Test und Validierung sowie eine k-fache Kreuzvalidierung[83] umgangen werden.[84]

Aufgrund des enormen Potenzials von ML werden auch bei der Prognose des BIP im Zuge der Digitalisierung anstelle von herkömmlichen Verfahren vermehrt algorithmierte Computersimulationen verwendet.[85] Auch der Sachver-

---

[78] Jung/Patnam/Ter-Martirosyan (2018), S. 4.

[79] Varian (2014), S. 6.

[80] Jung/Patnam/Ter-Martirosyan (2018), S. 7.

[81] Woloszko (2017), S. 7.

[82] Sachverständigenrat (2019), S. 54.

[83] Für eine genauere Erklärung der Kreuzvalidierung siehe Varian (2014).

[84] Varian (2014), S. 6–7.

[85] Jung/Patnam/Ter-Martirosyan (2018), S. 4.

ständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung weist in seinem Jahresgutachten 2019/20 darauf hin, dass Methoden des ML eine potenzielle Alternative zu herkömmlichen (BIP-)Prognoseverfahren darstellen. Dabei sieht er insbesondere in neuronalen Netzen eine gute Alternative zu den herkömmlichen Brückengleichungen. *Random Forests* wird hingegen eine niedrigere Prognosegüte zugeschrieben.[86]

Möglichkeiten und Grenzen von Nowcasts mit Big Data und maschinellem Lernen

Die Verwendung der oben genannten Methoden des maschinellen Lernens für Nowcasts mit Big Data eröffnet eine Vielzahl von Möglichkeiten für (Echtzeit-)Prognosen. Diese sind in der gegenwärtigen Literatur bisher nur wenig erforscht. Allerdings zeichnet sich das enorme Potenzial dieser Anwendung künstlicher Intelligenz immer mehr ab. Erste Ergebnisse bei der kurz- und langfristigen Prognose mit Methoden des maschinellen Lernens in verschiedenen Anwendungsbereichen finden sich in Campos Vázquez/López-Araiza B.[87], Jung/Patnam/Ter-Martirosyan[88] sowie bei der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD)[89].

Campos Vázquez/López-Araiza B.[90] erstellen sehr moderne Nowcasts eines anderen volkswirtschaftlichen Indikators mit Big Data und ML. Basierend auf Google-Trends-Daten prognostizieren sie mit Nowcasts die Arbeitslosigkeit in Mexiko. Dabei vergleichen sie den Vorhersagefehler verschiedener Modelle des ML. Den niedrigsten RMSE weist das LASSO-Verfahren auf (0,28). Es wurde mit kubischen Wechselwirkungen trainiert und kann sich deutlich besser an Situationen anpassen, in denen sich die Beziehung zwischen den Variablen ungewöhnlich verhält.[91] Das LASSO-Verfahren erzielt eine etwas höhere Prognosegüte als das herkömmliche autoregressive Modell. Des Weiteren verbessert die

---

[86] Sachverständigenrat (2019), S. 54–55.

[87] Campos Vázquez/López-Araiza B. (2020).

[88] Jung/Patnam/Ter-Martirosyan (2018).

[89] Woloszko (2017).

[90] Campos Vázquez/López-Araiza B. (2020).

[91] Campos Vázquez/López-Araiza B. (2020), S. 145–146.

Verwendung der Google-Trends[92]-Daten die Prognosen.[93] Der RMSE des AR-Modells mit Google Trends liegt bei 0,29 im Vergleich mit 0,31 ohne.[94]

Jung/Patnam/Ter-Martirosyan[95] testen die Prognosegüte von BIP-Forecasts für sechs verschiedene Länder für drei Methoden des ML: elastisches Netz, Super Lerner und rekurrentes neuronales Netzwerk. Dabei erstellen sie BIP-Prognosen für ein Quartal beziehungsweise ein Jahr im Voraus. Bei der Quartalsprognose ohne Weltwirtschaftskrisenzeitraum, von 2010 bis 2016, schneiden die ML-Verfahren elastisches Netz und Super Lerner besser ab als die Prognosen des Weltwirtschaftsausblicks (*world economic outlook* [WEO]) des Internationalen Währungsfonds (IWF). Für die Quartalsprognose mit Krisenzeitraum (2007 bis 2016), wird das jeweils beste ML-Modell je Land dem WEO-Modell, dem AR-Modell sowie dem VAR-Modell gegenübergestellt. Der RMSE ist für alle Länder bei Prognose mit ML am niedrigsten und liegt zwischen 0,20 und 0,81. Im Fall des VAR-Modells liegt der RMSE zwischen 1,09 und 2,20, im Fall des AR-Modells zwischen 0,80 und 1,52 und bezogen auf das WEO-Modell zwischen 1,11 und 2,20 – je nach Land. Die Prognose für ein Jahr im Voraus erzielt deutlich schlechtere Ergebnisse und ist nur teilweise besser als die IWF-Standardprognose.[96] Allgemein sind Jung/Patnam/Ter-Martirosyan der Überzeugung, dass mit ML erstellte Prognosen Expertenprognosen nicht ersetzen, sondern viel mehr als zusätzliche Hilfestellung für die finale Prognose dienen sollten.[97]

Die OECD erstellt Vorhersagen des BIP-Wachstums der G6-Länder[98] in Pseudo-Echtzeit mit adaptiven Bäumen. Die Prognosen anhand dieses Verfahrens liefern bessere Ergebnisse als Standardmodelle wie das OECD-Indikatormodell oder die Konsensprognose.[99] Werte des RMSE liegen für diese Methode leider nicht vor, sondern nur die Angabe, um wie viel Prozent der RMSE im Vergleich mit den Standardmodellen verbessert werden konnte. Die ermittelte

---

[92] Google Trends ist ein täglicher beziehungsweise monatlicher Echtzeitindex, der das Abfragevolumen von Google-Nutzern erfasst.

[93] Campos Vázquez/López-Araiza B. (2020), S. 147–148.

[94] Campos Vázquez/López-Araiza B. (2020), 146.

[95] Jung/Patnam/Ter-Martirosyan (2018).

[96] Jung/Patnam/Ter-Martirosyan (2018), S. 22–24.

[97] Jung/Patnam/Ter-Martirosyan (2018), S. 27.

[98] Großbritannien, USA, Frankreich, Japan, Deutschland und Italien Woloszko (2017), S. 25.

[99] Woloszko (2017), S. 27.

Verbesserung der Prognosegüte liegt bei 12 bis 23 % in Bezug auf das Indikatormodell und durchschnittlich bei 32 % für die Konsensprognose.[100]

Die beschriebenen Arbeiten deuten an, welch weitreichende Möglichkeiten sich für Nowcasts mit Big Data und ML bieten. Dennoch ergibt die Analyse der in diesem Beitrag beleuchteten Studien relativ große Schwankungen in der Prognosegüte der einzelnen Modelle. Dabei bleibt festzuhalten, dass es sich hierbei um keinen vollumfänglichen Vergleich der verschiedenen Verfahren handelt, da dieser Bereich noch nicht in ausreichendem Umfang untersucht wurde und die Studien aufgrund unterschiedlicher Datengrundlagen nur eingeschränkt vergleichbar sind. Stattdessen werden unterschiedliche Ansätze anhand von exemplarischen Studien einander gegenübergestellt, um einen ersten Eindruck von der Güte der Prognosen zu erhalten. Den höchsten Wert der mittleren Fehlerquadratsumme weisen in dieser Übersicht das VAR-Modell, ein AR-Modell sowie das Brückengleichungsmodell des GDPNow auf. Der RMSE dieser Verfahren kann oberhalb des Wertes zwei liegen. Deutlich niedrigere Werte ergeben das ML-Modell von Jung/Patnam/Ter-Martirosyan[101] sowie das Brückengleichungsmodell und das dynamische Faktormodell von Bańbura/Giannone/Modugno/Reichlin[102] mit Werten unter eins. Den niedrigsten RMSE erzielen Campos Vázquez/López-Araiza B.[103] mit der LASSO-Methode und mit der Random-Forest-Schätzung sowie Bańbura/Giannone/Reichlin[104] mit einem dynamischen Faktormodell. Alle drei RMSE sind niedriger als 0,3.

Die teilweise großen Unterschiede zwischen den jeweiligen RMSE innerhalb einer Studie lassen sich bei Vergleichen verschiedener Länder vermutlich auf die länderspezifischen Unterschiede hinsichtlich der Volatilität der Daten und der Zahl der Beobachtungen zurückführen. Bei mehreren Prognosen für das gleiche Land ergeben sich Unterschiede im RMSE durch das Hinzukommen von Informationen für die jeweilige prognostizierte Periode, beispielsweise das Quartal oder Jahr, im Zeitablauf, was die Prognose verbessert.

---

[100] Woloszko (2017), S. 23 und 25.

[101] Jung/Patnam/Ter-Martirosyan (2018).

[102] Bańbura et al. (2013).

[103] Campos Vázquez/López-Araiza B. (2020).

[104] Bańbura/Giannone/Reichlin (2011).

Tab. 1: RMSE verschiedener Prognosemodell

MODELL	RMSE	QUELLE
Autoregressives Modell	0,85 – 2,34	Cepni et al. (2020)
Autoregressives Modell	0,80 – 1,52*	Jung et al. (2018)
Autoregressives Modell	0,29 – 0,31	Campos Vázquez/López-Araiza B. (2020)
Vektorautoregressives Modell	1,09 – 2,20*	Jung et al. (2018)
Brückengleichungsmodell	0,58 – 0,66	Bańbura et al. (2013)
Brückengleichungsmodell (GDPNow)	1,15 – 2,2	Higgins (2014)
Dynamisches Faktormodell	0,18 – 0,28	Bańbura et al. (2011)
Dynamisches Faktormodell	0,43 – 0,57	Bańbura et al. (2013)
Dynamisches Faktormodell (ifoCAST)	0,32	Lehmann et al. (2020)
Bestes ML-Modell	0,20 – 0,81*	Jung et al. (2018)
LASSO	0,28	Campos Vázquez/López-Araiza B. (2020)
Random Forest	0,29	Campos Vázquez/López-Araiza B. (2020)

RMSE = root mean square error (Wurzel der mittleren Fehlerquadratsumme)
** Werte beziehen sich auf Quartalsprognose mit Krisenzeitraum (2007 bis 2016)*
Quelle: Eigene Darstellung.

Weitere potenzielle Frühindikatoren, die sich aus bisher kaum genutzten Datensätzen diverser Aktivitäten ableiten lassen, werden vom Statistischen Bundesamt untersucht. Beispielsweise kann mit Mobilfunkdaten die Mobilität der Bevölkerung dargestellt werden. Im Zuge der COVID-19-Pandemie wurden so allgemeine Unterschiede im Pendlerverhalten erkannt. Ein solcher Näherungswert kann als tagesaktueller Index der konjunkturellen Aktivität in bestimmten Gebieten verwendet werden. Als frühzeitige Umsatzprognose eignet sich die Auswertung von Umsatzsteuervoranmeldungen. Des Weiteren kann mit Daten zu Onlinetransaktionen sowie neuen Hypotheken- und Kreditverträgen die Aktivität im Onlinehandel, bei Hauskäufen oder beim privaten Konsum abgebildet werden.[105]

Zur frühzeitigen Prognose von Verbraucherpreisen können Scannerdaten aus dem Einzelhandel herangezogen werden. Sie enthalten Informationen über Art, Absatz und Umsatz verkaufter Artikel und werden an der Kasse erfasst. Dabei kann auch die Nachfrage nach bestimmten Produktgruppen analysiert werden. Außerdem kann man daraus die gegenwärtige wirtschaftliche Lage des Einzelhandels ableiten. Des Weiteren können Satellitendaten zur Bestimmung der wirtschaftlichen Aktivität herangezogen werden, zum Beispiel zum Aufkommen an Schiffen auf dem Meer oder von Autos vor einem Geschäft.[106] Als Frühindikator

[105] Hauf/Stehrenberg/Zwick (2020), S. 55–60.

[106] Hauf/Stehrenberg/Zwick (2020), S. 55–60.

für die Industrieproduktion in Deutschland kann ein an Arbeitstagen aktualisierter Lkw-Maut-Fahrleistungsindex verwendet werden. Die Fahrleistung großer Lkw kann damit mit fünf bis neuntägigem Abstand ziemlich genau erfasst werden.[107]

Eine weitere potenzielle Quelle von Big Data für ökonomische Analysen sind Textdaten. Ähnlich wie in den Beiträgen von Silke Sturm sowie von Deniz Dilan Karaman Örsal in diesem Band beschrieben, können Textdaten aus den Sozialen Medien automatisiert ausgewertet werden und als weiterer, sehr kurzfristiger Faktor in Nowcasts einfließen. Erste Erfahrungen im Bereich der Politikanalyse sowie bei der Prognose von Kaufverhalten deuten auf eine hohe Genauigkeit dieses Verfahrens hin. Manche Forscher:innen gehen sogar so weit zu sagen, dass Wörter die neuen Zahlen sind.[108] Mehr Informationen zur Verwendung von Textdaten aus Zeitungen für die Konjunkturanalyse mit Nowcasts finden sich auch in Thorsrud[109].

Mit der Nutzung von Big Data, den Methoden des ML sowie dem Nowcast selbst gehen einige Probleme einher. Beispielsweise verlangt die große Geschwindigkeit von Big Data eine Vielzahl notwendiger Reaktionen. Diese schnellen Reaktionen erfordern im Gegenzug enorme Ressourcen für die zeitnahe Sammlung, Speicherung und Verarbeitung der Daten. Außerdem können sowohl die Ungenauigkeit als auch die Verschiedenartigkeit der Daten und die damit einhergehende Komplexität teilweise nicht ohne Weiteres bewältigt werden.[110] So liegen viele Probleme mit Big Data in den hohen Anforderungen begründet, die diese Datenart an Systemarchitekturen, Datenverarbeitungsrahmen sowie Verarbeitungsmodi stellt.[111]

Außerdem kann es bei der Vielzahl von Datenpunkten leicht passieren, dass ein Ergebnis nur aufgrund der Menge der Daten statistisch signifikant ist und nicht aufgrund eines tatsächlichen Zusammenhangs. Darüber hinaus werden mit Big Data überwiegend Korrelationen und keine Kausalitäten aufgezeigt (Scheinkorrelationen). Auch das Auswerten von Daten ohne zugrunde liegende Theorie, die die Ergebnisse erklären könnte, kann zu Problemen führen.[112] Des-

[107] Cox et al. (2020).

[108] Thorsrud (2020), S. 393.

[109] Thorsrud (2020).

[110] Jin et al. (2015), S. 59 und 62.

[111] Jin et al. (2015), S. 63.

[112] Taylor/Schroeder/Meyer (2014), S. 6–7.

wegen bezweifeln Ökonom:innen häufig die Zuverlässigkeit und die Repräsentativität von Big Data.[113]

Aus unternehmerischer Sicht gibt es zwei weitere Herausforderungen. Einerseits existiert eine Kluft zwischen Unternehmen, die über viele Daten verfügen, zum Beispiel Social-Media-Plattformen, und solchen, die kaum Zugriff auf Daten haben. Andererseits müssen Unternehmen über entsprechende Technologien und Fertigkeiten im Umgang mit Big Data verfügen, um wettbewerbsfähig zu sein.[114]

Bezüglich des ML ist es problematisch, dass diese Verfahren oft nicht ohne Weiteres erklären können, welche Faktoren zu der Prognose geführt haben, weil der Analyse kein spezielles Modell zugrunde liegt. Deswegen sind die mit Methoden des ML erlangten Ergebnisse nur schwer nachvollziehbar und interpretierbar und bleiben häufig sozusagen in einer Black Box.[115]

Um langfristig als zuverlässige Konjunkturprognose anerkannt zu werden, müssen Nowcasts, die auf Big Data und ML-Methoden basieren, bestimmte Kriterien erfüllen. Laut Deutscher Bundesbank ist die Prognosegüte dabei der wichtigste Faktor, also die Frage, wie genau das jeweilige Modell eine relevante Zielvariable voraussagen kann. Außerdem sollte das Modell möglichst geringen Schwankungen unterliegen, um die Interpretation und Kommunikation aktueller Ergebnisse zu erleichtern. Darüber hinaus ist jedoch auch relevant, dass das Ergebnis mit Hilfe des Modells nachvollzogen werden kann.[116] Dies gestaltet sich, wie oben beschrieben, bei Verfahren des ML häufig schwierig. Abgesehen davon ist die frühzeitige Prognose von Trendwenden immens wichtig. Da dies nicht so leicht zu erreichen ist, sollte das Augenmerk auch auf die Vorhersage solcher Wendungen gelegt werden. Weitere Probleme beim ökonomischen Nowcasting, wie mögliche Messfehler, werden in Castle et al.[117] beleuchtet.

Überdies gilt es, auch ethische und datenschutzrechtliche Erwägungen zu berücksichtigen. Ungeachtet der Vorteile der Big-Data-Analyse kann es aus ethischer Sicht ungünstig sein, dass sowohl in der Gegenwart als auch in der Zukunft der freie Wille nicht geschützt ist, da mit großen Datenmengen gegen-

113 Taylor/Schroeder/Meyer (2014), S. 1.

114 Mauro/Greco/Grimaldi (2015), S. 100–101.

115 Sachverständigenrat (2019), S. 54; Jung/Patnam/Ter-Martirosyan (2018), S. 4–5.

116 Deutsche Bundesbank (2018), S. 17.

117 Castle/Fawcett/Hendry (2009).

wärtiges Verhalten untersucht und zukünftiges Verhalten prognostiziert und damit auch für manipulative Zwecke instrumentalisierbar wird.[118]

Außerdem ist aus datenschutzrechtlichen Gründen der Zugriff auf interessante Datensätze nur eingeschränkt möglich. Häufig werden Daten nur dort ausgewertet, wo sie gesammelt werden. Beispielsweise wird der Lkw-Maut-Fahrleistungsindex vom Bundesamt für Güterverkehr erstellt. Kooperationen mit anderen Einrichtungen wie dem Statistischen Bundesamt werden ebenfalls durchgeführt. Allerdings ist es nur schwer vorstellbar, dass solche Daten in Deutschland kurzfristig einer Vielzahl von Forscher:innen zur Verfügung gestellt werden. Auch Verwaltungsdaten sind hochinteressant für die Wirtschaftsforschung. Allerdings ist der Zugriff auf diese Daten je nach Land sehr umständlich.[119] Daten aus dem privaten Sektor können beispielsweise durch Kollaborationen mit privaten Firmen analysiert werden.[120] Abgesehen davon sind Verfügbarkeit und Zugriff auf Daten europaweit sehr unterschiedlich geregelt,[121] sodass Big-Data-Prognosen unter Berücksichtigung der jeweiligen länderspezifischen gesetzlichen Regelungen erfolgen müssen. Dennoch sind die mit solchen Datensammlungen machbaren Prognosen voller Potenzial und von großer Bedeutung für zukünftige Konjunkturvorhersagen.

Schlussfolgerungen

Konventionelle Methoden zur Vorhersage der Konjunkturentwicklung reichen nicht mehr aus, da sie Veränderungen und Wandel nicht genau genug abbilden und Trendwenden nicht frühzeitig genug erkennen. Wissenschaft und auch Gesellschaft verlangen nach neuen, besseren Prognoseverfahren. Dieser Beitrag hat gezeigt, dass es neue Möglichkeiten der Konjunkturprognose gibt, gerade auch durch, mit und dank der Datenökonomie.

Eine in der gegenwärtigen Literatur vieldiskutierte Alternative zu traditionellen Forecasts sind hochaktuelle Nowcasts, die, wenn sie auf hochfrequenten Daten beruhen, quasi in Echtzeit Informationen lesen und zu Prognosen der Gegenwart verarbeiten können. Verschiedene Verfahren verweisen auf die schier unendlichen Möglichkeiten der Kombination von Big Data und ML mit

[118] Mauro/Greco/Grimaldi (2015), S. 100.

[119] Einav/Levin (2014b), S. 2–3.

[120] Einav/Levin (2014b), S. 3.

[121] Hauf/Stehrenberg/Zwick (2020), S. 61.

Nowcasts. Im Allgemeinen zeichnet sich bei diesem Ansatz eine Tendenz zu erhöhter Prognosegüte gegenüber herkömmlichen Verfahren ab.

Momentan steht die Wissenschaft noch am Anfang dieser neuartigen Entwicklungen. Trotz der genannten Limitationen wird weiter an der Entwicklung von Nowcasts entscheidender makroökonomischer Variablen wie des BIP gearbeitet werden, da die Kenntnis der gegenwärtigen Konjunktur von großer Bedeutung ist, das Potenzial von Big Data enorm und ML für die effiziente Auswertung dieser Daten bestens geeignet ist. Der Drang nach Etablierung und Weiterentwicklung von Nowcasts zeigt sich beispielsweise an den enormen Fortschritten, die im vergangenen Jahr auf diesem Gebiet in Instituten und Zentralbanken gemacht wurden. Schon jetzt ermöglichen die verschiedenen Nowcasts einen ersten Schritt in Richtung Neuvermessung der Weltwirtschaft, indem Informationen zu konjunkturellen Aktivitäten (zum Beispiel das BIP) früher veröffentlicht werden, ohne große Einbußen an Genauigkeit hinnehmen zu müssen. Auf längere Sicht ist es denkbar, dass die Verfügbarkeit eines genauen, aktuellen Konjunkturindikators zu einem Wettbewerbsvorteil für Unternehmen, Institutionen und Länder werden kann. Wer früher als alle anderen über die gegenwärtigen Entwicklungen informiert ist, kann am besten auf die Wirtschaft einwirken und somit Tendenzen verstärken oder frühzeitig abfangen.

Die immense Bedeutung der riesigen Datenmengen (das heißt von Big Data), mit denen das Verhalten Einzelner, aber auch die Entwicklungen einer ganzen Volkswirtschaft relativ genau beschrieben werden können, wird immer deutlicher werden, auch für Nowcasts. Es zeigt sich bereits, dass die Erweiterung traditioneller Prognosen um diese Datenart die Genauigkeit von Vorhersagen erhöht. Um Big Data möglichst effizient auszuwerten, sollten insbesondere Methoden des ML zum Einsatz kommen und mit den Ergebnissen herkömmlicher Verfahren verglichen werden. Vieles deutet darauf hin, dass künstliche Intelligenz die Zukunft mitgestalten wird. Deswegen müsste auch die Wissenschaft von dieser Art, Informationen zu verarbeiten, Gebrauch machen. Dabei sollte ML nicht vorwiegend als Konkurrenz zu etablierten Methoden betrachtet werden, sondern als Möglichkeit, bestehende Ansätze zu erweitern beziehungsweise zu vertiefen.

Damit es nicht zu ethischen Problemen kommt, muss die Auswertung von Big Data sehr vorsichtig und nach international festgelegten Regeln und Gesetzen erfolgen. Dann bieten sich der Echtzeitprognose wichtiger Kennzahlen unzählige Möglichkeiten und die neue Datenökonomie kann selbst dazu beitragen, die mit ihr einhergehenden Herausforderungen zu bewältigen.

Literatur

Aastveit, K. A.; Gerdrup, K. R.; Jore, A. S.; Thorsrud, L. A. (2014): Now-casting GDP in Real Time: A Density Combination Approach. In: Journal of Business & Economic Statistics Vol. 32 (1), S. 48–68, DOI https://doi.org/10.1080/07350015.2013.844155.

An der Heiden, M.; Hamouda, O. (2020): Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting. In: Epidemiologisches Bulletin Vol. 17, S. 10–16, DOI https://doi.org/10.25646/6692.4.

Armstrong, J. S. (Hrsg.) (2004): Principles of Forecasting. A Handbook for Researchers and Practitioners. 4. Aufl., Kluwer Academic Publishers. New York, Boston, Dordrecht, London, Moskau (30).

Bańbura, M.; Giannone, D.; Modugno, M.; Reichlin, L. (2013): Now-casting and the Real-Time Data Flow. In: Elliott, G. und Timmermann, A. (Hrsg.): Handbook of Economic Forecasting, North Holland, Amsterdam, S. 195–237.

Bańbura, M.; Giannone, D.; Reichlin, L. (2011): Nowcasting. In: Clements, M. P. und Hendry, D. F. (Hrsg.): The Oxford Handbook of Economic Fore-casting, Oxford University Press, Oxford, New York, Auckland, S. 193–223.

Bok, B.; Caratelli, D.; Giannone, D.; Sbordone, A. M.; Tambalotti, A. (2018): Macroeconomic Nowcasting and Forecasting with Big Data. In: Annual Review of Economics Vol. 10 (1), S. 615–643, DOI https://doi.org/10.1146/annurev-economics-080217-053214.

Bundesministerium für Wirtschaft und Energie (2019): Nowcasting: Ein Echtzeit-Indikator für die Konjunkturanalyse. In: Schlaglichter der Wirtschaftspolitik (a), S. 12–15.

Bureau of Economic Analysis (2020): National Data. National Income and Product Accounts. Online verfügbar unter https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2#reqid=19&step=2&isuri=1&1921=survey, zuletzt geprüft am 05.11.2020.

Campos Vázquez, R. M.; López-Araiza B., S. E. (2020): Grandes Datos, Google y Desempleo - Big Data, Google and Unemployment. In: Estudios Económicos Vol. 35 (1), S. 125-151, DOI https://doi.org/10.2307/26863997.

Castle, J. L.; Fawcett, N. W.; Hendry, D. F. (2009): Nowcasting is not Just Contemporaneous Forecasting. In: National Institute Economic Review Vol. 210, S. 71–89, DOI https://doi.org/10.1177/0027950109354412.

Cepni, O.; Guney, I. E.; Swanson, N. R. (2020): Forecasting and Nowcasting Emerging Market GDP Growth Rates. The Role of Latent Global Economic Policy Uncertainty and Macroeconomic Data Surprise Factors. In: Journal of Forecasting Vol. 39 (1), S. 18–36, DOI https://doi.org/10.1002/for.2602.

Cox, M.; Triebel, J.; Linz, S.; Fries, C.; Flores, L. F.; Lorenz, A. et al. (2020): Täglicher Lkw-Maut-Fahrleistungsindex aus digitalen Prozessdaten der Lkw-Mauterhebung. In: WISTA - Wirtschaft und Statistik Vol. 72 (4), S. 63–76.

Deutsche Bundesbank (2018): Modelle zur kurzfristigen Konjunkturprognose: eine Aktualisierung. In: Monatsbericht, September, S. 15–29.

Deutsche Bundesbank (2020): Methodik. Wöchentlicher Aktivitätsindex. Online verfügbar unter https://www.bundesbank.de/de/statistiken/konjunktur-und-preise/woechentlicher-aktivitaetsindex/methodik-833778, zuletzt geprüft am 25.9.20.

Deutscher Wetterdienst (DWD) (2020): Nowcasting-Verfahren. Online verfügbar unter https://www.dwd.de/DE/forschung/wettervorhersage/met_fachverfahren/nowcasting/nowcasting_node.html, zuletzt geprüft am 28.7.2020.

Dickopf, X.; Janz, C.; Mucha, T. (2019): Vom BIP-Flash zum BIP-Nowcast. Erste Ergebnisse einer Machbarkeitsstudie zur weiteren Beschleunigung der BIP-Schnellschätzung. In: WISTA - Wirtschaft und Statistik Vol. 71 (6), S. 47-58.

Doody, R.; Kokkinen, A.; Wouters, H. (2016): Euro Area and European Union GDP Flash Estimates at 30 Days, Publications Office. Luxemburg.

Einav, L.; Levin, J. (2014): Economics in the Age of Big Data. In: Science Vol. 346 (6210), S. 1–6, DOI https://doi.org/10.1126/science.1243089.

Eurostat (2019): Vorläufige Schnellschätzung des BIP für Europa 30 Tage nach Quartalsende. Online verfügbar unter https://ec.europa.eu/eurostat/statistics-explained/index.php/Preliminary_GDP_flash_estimate_in_30_days_for_Europe/de, zuletzt aktualisiert am 19.8.19, zuletzt geprüft am 25.9.20.

Federal Reserve Bank of Atlanta (2020): GDPNow. Online verfügbar unter https://www.frbatlanta.org/cqer/research/gdpnow?panel=2, zuletzt geprüft am 25.9.20.

Federal Reserve Bank of New York (2019): Nowcasting and Forecasting with Big Data. Unter Mitarbeit von Time Series Analysis Team. Federal Reserve Bank of New York.

Federal Reserve Bank of New York (2020): Nowcasting Report. Online verfügbar unter https://www.newyorkfed.org/research/policy/nowcast.html, zuletzt geprüft am 19.11.2020.

Garboden, P. M. E. (2020): Sources and Types of Big Data for Macroeconomic Forecasting. In: Fuleky, P. (Hrsg.): Macroeconomic Forecasting in the Era of Big Data. Theory and Practice. 1. Aufl., Springer International Publishing, Cham, S. 3–23.

Giannone, D.; Reichlin, L.; Small, D. (2008): Nowcasting: The real-time informational content of macroeconomic data. In: Journal of Monetary Economics Vol. 55 (4), S. 665–676, DOI https://doi.org/10.1016/j.jmoneco.2008.05.010.

Günther, F.; Bender, A.; Katz, K.; Küchenhoff, H.; Höhle, M. (2020): Now-casting the COVID-19 Pandemic in Bavaria. Ludwig-Maximilians-Universität München.
Gupta, M.; Minai, M. H. (2019): An Empirical Analysis of Forecast Performance of the GDP Growth in India. In: Global Business Review Vol. 20 (2), S. 368–386, DOI https://doi.org/10.1177/0972150918825207.
Hauf, S.; Stehrenberg, S.; Zwick, M. (2020): EXDAT - Experimentelle Daten und Methoden für eine Innovative Statistik. In: WISTA - Wirtschaft und Statistik Vol. 72 (4), S. 51–62.
Higgins, P. (2014): GDPNow: A Model for GDP „Nowcasting". In: Federal Reserve Bank of Atlanta Working Paper Series Vol. 7.
Höhle, M.; An der Heiden, M. (2014): Bayesian nowcasting during the STEC O104:H4 outbreak in Germany, 2011. In: Biometrics Vol. 70 (4), S. 993–1002, DOI https://doi.org/10.1111/biom.12194.
Jin, X.; Wah, B. W.; Cheng, X.; Wang, Y. (2015): Significance and Challenges of Big Data Research. In: Big Data Research Vol. 2 (2), S. 59–64, DOI https://doi.org/10.1016/j.bdr.2015.01.006.
Jung, J.-K.; Patnam, M.; Ter-Martirosyan, A. (2018): An Algorithmic Crystal Ball. Forecasts-based on Machine Learning. Washington, D.C. (IMF working paper).
Lawless, J. F. (1994): Adjustments for reporting delays and the prediction of occurred but not reported events. In: The Canadian Journal of Statistics Vol. 22 (1), S. 15–31.
Lehmann, R.; Reif, M.; Wollmershäuser, T. (2020): ifoCAST: Der neue Prognosestandard des ifo Instituts. In: ifo Schnelldienst Vol. 73 (11), S. 31–39. Online verfügbar unter https://www.ifo.de/publikationen/2020/aufsatz-zeitschrift/ifocast-der-neue-prognosestandard-des-ifo-instituts, zuletzt geprüft am 1.7.2021.
Mauro, A. de; Greco, M.; Grimaldi, M. (2015): What is big data? A consensual definition and a review of key research topics. In: AIP Conference Proceedings Vol. 1644 (1), S. 97–104, DOI https://doi.org/10.1063/1.4907823.
Pons, J. (2000): The accuracy of IMF and OECD forecasts for G7 countries. In: Journal of Forecasting Vol. 19 (1), S. 53–63.
Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2019): Den Strukturwandel Meistern. Jahresgutachten 19/20.
Statistisches Bundesamt (2020): Volkswirtschaftliche Gesamtrechnung. BIP Schnellschätzung t+30. Online verfügbar unter https://www.destatis.de/DE/Themen/Wirtschaft/Volkswirtschaftliche-Gesamtrechnungen-Inlandsprodukt/AnkuendigungBIPFlash.html, zuletzt aktualisiert am 30.7.2020, zuletzt geprüft am 3.12.2020.

Stock, J. H.; Watson, M. W. (2017): Twenty Years of Time Series Econometrics in Ten Pictures. In: Journal of Economic Perspectives Vol. 31 (2), S. 59–86, DOI https://doi.org/10.1257/jep.31.2.59.

Taylor, L.; Schroeder, R.; Meyer, E. (2014): Emerging Practices and Perspectives on Big Data Analysis in Economics: Bigger and Better or More of the Same? In: Big Data & Society Vol. 1 (2), S. 1-10, DOI https://doi.org/10.1177/2053951714536877.

Thorsrud, L. A. (2020): Words are the New Numbers: A Newsy Coincident Index of the Business Cycle. In: Journal of Business & Economic Statistics Vol. 38 (2), S. 393–409, DOI https://doi.org/10.1080/07350015.2018.1506344.

Varian, H. R. (2014): Big Data: New Tricks for Econometrics. In: Journal of Economic Perspectives Vol. 28 (2), S. 3–28, DOI https://doi.org/10.1257/jep.28.2.3.

Vosen, S.; Schmidt, T. (2011): Forecasting Private Consumption: Survey-based Indicators vs. Google Trends. In: Journal of Forecasting Vol. 30 (6), S. 565–578, DOI https://doi.org/10.1002/for.1213.

Woloszko, N. (2017): Economic Modelling and Machine Learning. A Proof of Concept. Konferenzpapier. Organisation for Economic Co-operation and Development. Paris.

World Meteorological Organization (2017): Guidelines for Nowcasting Techniques.

Textdaten

Anwendungen und Herausforderungen

Silke Sturm

Einleitung

Textdaten sind die am stärksten wachsende Datenquelle. Seit zwei Jahrzehnten nimmt die Zahl der Haushalte mit dauerhaftem Internetzugang zu, häufig ist dieser durch mobile Daten zusätzlich unabhängig von ihrem Aufenthaltsort.[1] Es werden auf Seiten aller gesellschaftlichen und wirtschaftlichen Akteure Beiträge verfasst, wobei sowohl redaktionell bearbeitete als auch privat verfasste Texte relevant sind. Durch die größere Bedeutung von Textdaten und die sich stetig verbessernde Rechenleistung entwickelt sich der Bereich automatisierter Textanalysen in vielfältigen Fachbereichen dynamisch.

Die Nutzung einer großen Bandbreite von Veröffentlichungen hat Vorteile in Form einer allgemeineren Abdeckung relevanter Themen und Meinungen und damit einer besseren Abschätzbarkeit von Entwicklungen. Allerdings ist durch die Varianz der Kommunikation, welche sich durch Länge, Qualität, Vokabular oder Menge der verschiedenen Themen unterscheidet, die Auswertung schwierig und bedarf eines präzisen Vorgehens. Das Ziel, die Chancen der Auswertung und das Verständnis wirtschaftlicher und politischer Zusammenhänge zu verbessern, ist die treibende Kraft hinter den aktuell steigenden Forschungsanstrengungen.[2]

Dieser Beitrag beschäftigt sich mit der automatisierten Textanalyse und ihren Vor- und Nachteilen. Darüber hinaus widmet er sich einer Betrachtung der Chancen und Herausforderungen Sozialer Medien und politisch-gesellschaftli-

---

[1] Kumar/Das (2013); WorldBank (2019).

[2] Einav/Levin (2014).

cher Texte. Die Anwendbarkeit unstrukturierter Textdaten für Forschungsinteressen wird anhand eines exemplarischen Auszugs aus der politischen Diskussion in der Legislaturperiode 2014 bis 2017 dargestellt.

Texte als Daten

Vor- und Nachteile textbasierter Daten

Der Vorteil strukturierter Daten in Form von Statistiken, Surveys, tabellarischen Aufstellungen oder weiteren quantitativen Erhebungen ist naheliegend. Die Daten liegen bereits in maschinenlesbarer Form vor, können ohne umfangreiche Vorbereitungen genutzt werden und sind weitestgehend vergleichbar. Statistiken ökonomisch und gesellschaftlich relevanter Daten liegen zum Teil in Aggregaten oder mit Zeitverzug vor. Insbesondere makroökonomische Variablen wie zum Beispiel Wachstum, Inflation, Konsum oder Arbeitslosigkeit stehen erst mit zeitlichem Verzug, also als Spätindikatoren zur Verfügung.[3] Es können demnach erst ex-post Veränderungen beziehungsweise vergangene Entwicklungen beobachtet werden. Zur Prognose zukünftiger Entwicklungen werden Forecasts über vorauslaufende Indikatoren wie zum Beispiel Börsenkurs, Bauaktivitäten oder Kredite genutzt.[4]

Um Verzögerungen auszugleichen, werden, etwa im Bereich des Konsumentenverhaltens, aufwendige Befragungen durchgeführt. Diese geben teilweise einen Ex-ante-Eindruck der Verhaltensentwicklung, sind dabei allerdings sehr aufwendig.[5]

Strukturierte, quantitative Daten bilden nur einen kleinen Teil der Realität und der tatsächlich zur Verfügung stehenden Daten ab. Durch die allgemeine Verfügbarkeit des Internets wird davon ausgegangen, dass weltweit bis zu 90 % der generierten Daten in unstrukturierter Form vorliegen.[6] Diese Daten umfassen Text- und Bilddaten aus den verschiedensten Zusammenhängen. Dabei sind Text- und Bilddaten nicht nur für makroökonomische Prognosen relevant, sondern fallen auch in politischen, gesellschaftlichen und unternehmerischen

---

[3] Iyetomi et al. (2020).

[4] Conference Board.

[5] OECD (2017); GfK.

[6] Einav/Levin (2014).

Kontexten in großen Mengen an. Big Data stellt eine große Chance für die Forschung und politische Entscheidungsträger dar, stellt sie aber auch vor die Herausforderung, die Daten aufwendig zu analysieren, Datenschutzrichtlinien zu beachten und eine Reproduzierbarkeit der Ergebnisse sicherzustellen.

Textdaten sind insbesondere auch in der ökonomischen Forschung von Bedeutung. Die Daten werden durch verschiedene Akteure in gesellschaftlichen, sozialen und geschäftlichen Beziehungen generiert. Historische Daten sind vornehmlich offizieller oder redaktioneller Natur, darunter fallen offizielle Reden, journalistische Beiträge, Pressemitteilungen oder Interviews. Durch Soziale Netzwerke, persönliche Blogs oder Websites und Kommentarspalten bilden informelle Informationen einen wachsenden Anteil an der gesamten Datenbasis. Zusätzlich haben sich die Menge der Informationen über einzelne Individuen und deren Aussagekraft für ökonomische und (wirtschafts-)politische Kontexte erhöht.[7]

Zusammengefasst bestehen drei elementare Vorteile der Verwendung textbasierter Daten:

- Die Daten sind ad hoc verfügbar. Es besteht kein Zeitverzug zwischen der Generierung der Daten und ihrer Verwertbarkeit.[8]
- Soziale Medien wie Twitter und Facebook oder Google-Anfragen stellen oft Meinungsäußerungen bewusster oder unbewusster Natur dar.[9]
- Textbasierte Daten sind kostengünstiger als Surveys.

Die zeitgenaue Verfügbarkeit der Daten ist im Vergleich mit traditionellen Statistiken als einer der größten Vorteile zu betrachten. Während für die Verarbeitung und Berechnung von Indizes Daten aus verschiedenen Quellen verarbeitet werden müssen und zum Teil erst verspätet zur Verfügung gestellt werden, sind Textdaten mit ihrer Veröffentlichung rascher verfügbar. Je nach Analysemethode können die Daten damit zeitnah Aufschluss über Veränderungen der Meinung oder der ökonomischen Aktivität liefern. Zusätzlich sind durch Mitteilungen in den Sozialen Medien unmittelbare Äußerungen der Akteure verfügbar. Über unmittelbare Meinungsäußerungen können Entwicklungen nicht nur zeitnah nachverfolgt, sondern auch vorhergesagt werden. Intensiv hat sich die Forschung unter anderem zur Entwicklung des Finanzmarkts nach Äußerun-

---

[7] Hu/Liu (2012).

[8] Ebenda; Sakaki/Okazaki/Matsuo (2010).

[9] Ravi/Ravi (2015).

gen der Zentralbanken entwickelt und kommt zu sehr guten Ergebnissen. Der Forschungsstrang nutzt Sentiments, um zum Beispiel Entwicklungen auf dem Aktienmarkt zu prognostizieren.[10] Diese Methodik kann entsprechend auch auf Mitteilungen von Konsumenten/Parteien/Wählern angewandt werden, um Verhalten und Entwicklungen vorauszusagen. Die Analysemethoden erlauben zusätzlich die Analyse impliziter Inhalte wie zum Beispiel psychologischen Motiven oder Sentiments. Durch die Nutzung impliziter Inhalte werden die unmittelbaren Äußerungen von in-(offiziellen) Akteuren noch besser interpretierbar.

Die Verwendung von Textdaten erfordert umfangreiche Vorbereitungen der Dokumente, die Vorteile der Textdaten müssen den Kosten und dem Aufwand der Auswertung gegenübergestellt werden. Als Dokumente/Textabschnitte zählen dabei einzelne abgeschlossene Texteinheiten, die für die spätere Analysemethode verwendet werden. Je nach verwandter Methodik können dies gesamte Dokumente, Sätze oder Wortgruppen sein. Die Vorbereitung der Textdaten ist wesentlich umfangreicher und zeitaufwendiger als die Vorbereitung von Daten im Rahmen offizieller Statistiken.[11]

Ein weiterer kritischer Aspekt der computerbasierten Textanalyse ist die Reproduzierbarkeit der Ergebnisse. Computerbasierte Textanalyse greift auf Algorithmen mit festgelegten Parametern zurück oder kodiert die Daten innerhalb von überwachten Ansätzen. Die entsprechenden Parameter müssen klar festgehalten und dargestellt werden, um eine Reproduktion zu ermöglichen. Für die händische Kodierung sind klare Richtlinien zur Validität und zur Reliabilität festgeschrieben, unüberwachte Verfahren erfordere eine ähnliche Vorgehensweise. Die Handkodierung eines Subsets der Dokumente kann zum Beispiel dazu beitragen, die Ergebnisse zu validieren.

Es gibt eine wachsende Zahl von Algorithmen zur Auswertung unstrukturierter Textdaten. Eine Erfassung aller Algorithmen ist nicht möglich, sie verteilen sich auf verschiedene Fachbereiche. Die Algorithmen koexistieren und werden je nach Anwendung und Fachbereich eingesetzt. Die Menge der Algorithmen wächst mit den Anwendungsfeldern. Grundsätzlich wird zwischen überwachten und unüberwachten Analysemethoden unterschieden, hinzu kommen Methoden wie Regressionsanalysen, Inhaltsanalysen oder Natural Language Processing (NLP).

[10] Tetlock (2007); Bollen/Mao/Zeng (2011).

[11] Gentzkow/Kelly/Taddy (2019).

Da in der ökonomischen Literatur die letztgenannten Methoden nur geringe Anwendung finden, sind sie für den vorliegenden Artikel von untergeordneter Relevanz. Quinn et al.[12] haben eine Liste der Kosten und der benötigten Datengrundlagen erstellt. Mit Ausnahme des Lesens von Texten und des automatisierten Topic Modelings setzen alle Methoden eine sehr gute Kenntnis der zugrundeliegenden Texte voraus. Die Kosten der Präanalysephase sind somit hoch. Die Kodierung der Texte beziehungsweise die Erstellung eines passgenauen Wörterbuchs oder die vollständig händische Kodierung erfordern die Bearbeitung eines Anteils oder des gesamten Textbestands. In Abbildung 1 werden die wichtigsten Algorithmenklassen für die automatisierte Textanalyse dargestellt. Einige ausgewählte Algorithmen werden in der Folge genauer vorgestellt.[13]

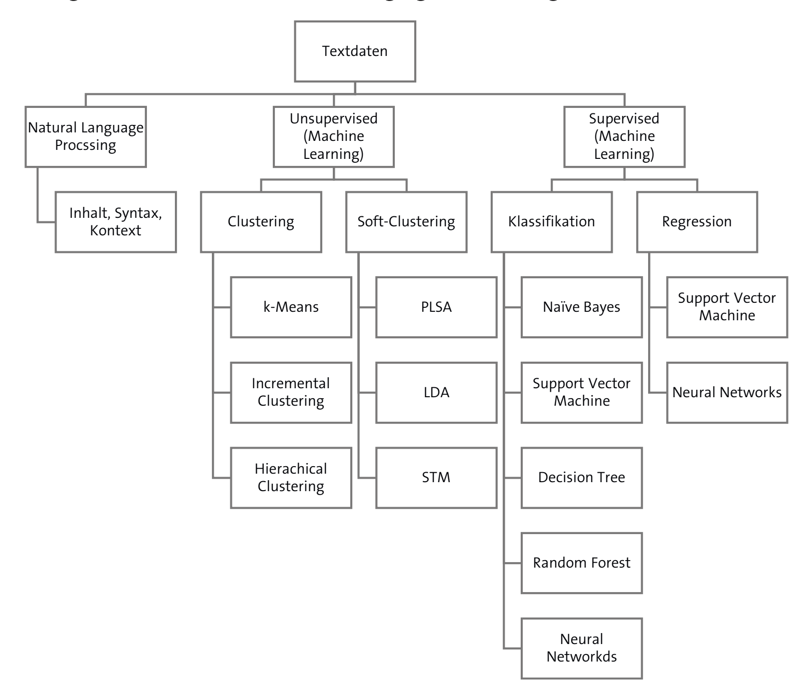


Abb. 1: Schematische Darstellung automatisierten Text Minings
Quelle: eigene Darstellung.

[12] Quinn et al. (2010).

[13] Für eine umfassende Darstellung bieten sich zum Beispiel Murphy (2012), Bishop (2016) und Gentzkow/Kelly/Taddy (2019) an.

Merkmalsextraktion

Für die Auswertung textbasierter Daten ist eine Reduktion der Dimension notwendig. Hierfür werden durch die Filterung unwichtiger Elemente charakteristische Merkmale herausgefunden. Die Extraktion von Merkmalen (*feature extraction*) wird zum Beispiel durch Löschen der Interpunktion erzielt. Zusätzlich werden häufig sogenannte Stop Words aus dem Datensatz entfernt. Stop Words sind üblicherweise häufig genutzte Wörter einer Sprache, wie zum Beispiel die Füllwörter „und", „weil" oder „da". Diese Wörter haben für die Extraktion des Inhaltes kaum Bedeutung, in spezifischen Kontexten kommen weitere Wörter hinzu. Dies können etwa Wörter sein, die zwar nur einer der zu kategorisierenden Gruppen zuzuordnen sind, jedoch nicht zur Beantwortung der Forschungsfrage dienlich sind. Beispielsweise können Politikernamen spezifisch für eine Partei sein, für die Frage des Parteienfokus sind sie jedoch nicht dienlich. Die Nutzung von Stop Words muss für jedes Studiendesign individuell geklärt werden.[14]

Eine Alternative ist die Verwendung von *term frequency – inverse document frequency-(tf-idf-)*Verfahren. Damit gemeint ist die Häufigkeit eines Wortes in einem Dokument, multipliziert mit dem logarithmierten Anteil an Dokumenten, die das Wort enthalten. Im Ergebnis wird ein niedriger tf-idf-Wert für sehr häufige Wörter in allen Dokumenten und sehr seltene Wörter in einem spezifischen Dokument ausgegeben. Anders formuliert werden als Features jene Wörter behalten, die häufig in einem einzelnen Dokument vorkommen, aber selten in anderen Dokumenten.[15]

Als weiterer Schritt wird ein Stemming, das heißt eine Verkürzung der Wörter auf ihren Wortstamm, durchgeführt. Alle beschriebenen Schritte reduzieren die Menge der zu analysierenden Daten. Zwei Aspekte profitieren von den Vorbereitungen des Datensatzes: Zum einen werden die Ergebnisse leichter interpretierbar und zum anderen sinkt die notwendige Rechenleistung.

---

[14] Nenkova/McKeown (2012).

[15] Ebenda.

Überwachte Verfahren

Überwachte oder Klassifizierungsmethoden bringen auf Grundlage eines Trainingsdatensets dem Modell Kriterien bei, die später für unbekannte Daten genutzt werden können, um Vorhersagen über Inhalte oder Verhaltensweisen zu treffen. Dem Trainingsdatenset werden dabei Kategorisierungen zugeordnet, der Algorithmus lernt daraufhin die Charakteristika der Kategorien und ordnet sie den Kategorien zu. In einfachen Ansätzen werden hierzu etwa Kritiken mit einer quantifizierbaren Bewertung als Kategorisierung genutzt. Es ist demnach keine weitere Kodierung durch Experten notwendig (zum Beispiel Produktbewertungen)[16].

Überwachte Methoden gelten insbesondere bei tiefgehenden inhaltlichen sowie psychologischen Textanalysen als präziser. Die Aussagen können vor dem automatisierten Verfahren ähnlich wie bei händischen Verfahren kodiert werden. Durch die Vorkodierung anhand theoriegeleiteter Kriterien verringert sich die Wahrscheinlichkeit einer fehlerhaften Zuordnung. Bei der Auswahl der Algorithmen sind zwei Aspekte zu beachten: Zum einen ist ein sogenanntes Overfitting des Modells möglich. Wenn der Algorithmus das Trainingsdatenset zu kleinteilig analysiert hat, kann es passieren, dass das Ergebnis auf der Basis neuer Daten sich schlechter darstellt. Die detaillierten Charakteristika des Trainingsdatensets müssen nicht im neuen Datensatz vorkommen und können damit nicht vom Algorithmus erfasst werden.[17] Zum anderen werden für überwachte Methoden mehr Daten benötigt, das heißt, es werden ein ausreichend großes, passgenaues Trainingsdatenset und die eigentlich zu untersuchenden Daten benötigt.[18]

Die Trainingsmethode basiert bei den meisten Studien auf Naive Bayes, Entscheidungsbäumen, Support Vector Machines oder regressionsbasierten Klassifizierungen.[19]

---

[16] Liu/Zhang (2012).

[17] Murphy (2012).

[18] Aggarwal (2012).

[19] Für einen umfangreichen Überblick bieten sich Gentzkow/Kelly/Taddy (2019) und Aggarwal/Zhai (2012a) an.

Unüberwachte Verfahren

In Bezug auf unüberwachte Verfahren wird zwischen klassischen Clustering-Ansätzen und dem sogenanntem Soft Clustering unterschieden. Anders als im Falle der überwachten Verfahren benötigen die Algorithmen keine vorkodierten Trainingsdaten. Aus den zur Verfügung stehenden Daten werden Cluster generiert, welche die Struktur der Texte widerspiegeln. Mit dem Ansatz können Informationen aus großen Mengen textbasierter Daten extrahiert werden. Grundsätzlich finden zwei Ansätze Verwendung: das klassisches Clustering und das Soft Clustering.

Beim klassischen Clustering wird die Anzahl der Cluster k definiert und die k Punkte werden randomisiert als Zentrum des Clusters festgelegt. Alle Bestandteile werden dem nächstliegenden Cluster-Zentrum zugeordnet (e*uclidean distance*). Anschließend wird der Mittelwert jedes Clusters berechnet. Dieser bildet sodann das Zentrum der Cluster. Die Methode ist schnell und effizient, kann jedoch etwa durch die präzise Wahl des initialen Zentrums erheblich verbessert werden. Des Weiteren ist die Anzahl der k Cluster nicht immer leicht zu bestimmen.[20]

Im Gegensatz zum deterministischen Ansatz des einfachen Clustering liegt dem probabilistischen Ansatz eine Dichteverteilung zugrunde. Soft-Clustering-Methoden (oder „Topic-Modell“) errechnen die Wahrscheinlichkeit, mit der ein Dokument einem Themencluster zugeordnet werden kann. Ein erster Versuch, probabilistische Aspekte in die Analyse zu integrieren, war die „probabilistic Latent Sementic Analysis“ (pLSA).[21] Dieser Methode liegt kein generatives Modell zugrunde, doch die Latent Dirichlet Allocation (LDA) eliminiert dieses Problem. Der Auswertung in diesem Beitrag liegt die LDA zugrunde.[22] LDA nimmt an, dass sich jedes Dokument aus einer Mischung von Themen zusammensetzt. Jedes Thema besteht aus einer Verteilung des über den Datensatz vorhandenen Vokabulars. Der probabilistische, generative Prozess ist wie folgt definiert:

---

[20] Aggarwal/Zhai (2012b).

[21] Hofmann (2013).

[22] Blei/Ng/Jordan (2003).

Für jedes Thema:

$$\beta k \sim DirV\,(\eta) \tag{1}$$

Für jedes Dokument:

$$\theta_d \sim Dir(\alpha) \tag{2}$$

Für jedes Wort:

$$Z_{d,n} \sim \mathrm{Mult}(\theta_d), \text{with } Z_{d,n} \in \{1, \dots, K\} \tag{3}$$

$$W_{d,n} \sim \mathrm{Mult}\left(\beta_{Z_{d,n}}\right), \text{with } W_{d,n} \in \{1, \dots, V\}$$

Die LDA verwendet zwei Dirichlet-Zufallsvariablen: erstens die Themen β_k, welche eine Verteilung über das Vokabular V (mit den Themen 1 bis K) sind, und zweitens die Themenzuordnung pro Dokument θ_d. Die Dirichlet-Verteilung verwendet feste Parameter η und α, die die Themenverteilung über Wörter beziehungsweise die Themenverteilung pro Dokument über Wörter beeinflussen. Die Themenzuweisung pro Wort $Z_{d,n}$ ist die dritte verborgene Variable, welche die zugrunde liegende latente Struktur des Korpus definiert. $W_{d,n}$ ist die beobachtete Variable, die Informationen über das im Korpus verwendete Vokabular enthält.[23] In anderen Worten liefert der Algorithmus eine Zusammenfassung der im Korpus enthaltenen Themen, bestehend aus dem spezifischen Vokabular und der Wahrscheinlichkeit, dass ein Thema einem Dokument zugeordnet werden kann. Seit den ersten Topic-Modells ist eine Vielzahl neuer Modelle mit unterschiedlicher Schwerpunktsetzung, zum Beispiel Themenkorrelation[24], strukturelle Topic-Modelle (STM)[25] oder dynamische Topic-Modelle[26], entstanden.

Politische und gesellschaftliche Texte

Direkte politische oder gesellschaftliche Kommunikation unterscheidet sich durch zwei Merkmale von formalen oder journalistischen Texten. Zum einen beinhalten die schnell verfassten, nicht redaktionell bearbeiteten Kommentare

---

[23] Blei/Lafferty (2009).

[24] Blei/Lafferty (2007).

[25] Roberts et al. (2014).

[26] Blei/Lafferty (2006).

ein vielseitiges Vokabular. Zum anderen steht zwischen dem Verfasser der Texte und den Lesenden kein Interpretationsschritt durch Erläuterungen, Zusammenfassungen oder journalistische Aufarbeitung.[27]

Im Allgemeinen unterscheiden sich Kommentare von politischen Akteur:innen und User:innen in vielerlei Hinsicht von anderen für inhaltsanalytische Ansätze genutzten Texten. Die Texte beinhalten eine wenig formalisierte Sprache mit einem großen und diversen Vokabular, zusätzlich erschweren Rechtschreibungsfehler oder Abkürzungen die Analysen.[28] Des Weiteren sind die Themen vielschichtig, das heißt, es wird sowohl innerhalb einer Nachricht als auch in der Grundgesamtheit die Bandbreite der gesellschaftlichen Kommunikation abgedeckt. Es gibt keine Kodierung dieser Themenbandbreite, sie wird in unkodierter und unstrukturierter Form veröffentlicht. Diese Charakteristika erschweren eine überwachte Analyse dieser Kommunikation. Der Diskurs umfasst eine stetig steigende Menge an Nachrichten. Zudem verändert sich der thematische Schwerpunkt im Zeitverlauf. Während manche Themen an Bedeutung verlieren oder verschwinden, werden andere Themen und damit auch ein erweitertes Vokabular relevant. Sichtbar ist die Wandlung von Themenschwerpunkten ex post, zum Beispiel in handkodierten Parteiprogrammen.[29]

Das Themenspektrum deckt zeitlich und gesellschaftlich die vorhandene Bandbreite ab. Diese Vielseitigkeit legt die Relevanz der Texte nahe. Es werden bevölkerungsrelevante Themen in nicht interpretierter oder gefilterter Form analysiert. Es können Themen erkannt werden, die in Surveys oder anderweitigen theoriegeleiteten Analysemethoden noch nicht abgefragt oder erkannt werden können. Die datengetriebene, explorative Methodik hat demgegenüber den Vorteil, dass sie die vorhandene Kommunikation in umfassender Form betrachtet und auswertet.

Anwendungsbereiche können unter anderem im politischen Kontext (Parteienkommunikation) und im Konsument:innenverhalten zu finden sein. Im politischen Kontext können sowohl spezifische wirtschaftspolitische Themen erkannt werden, als auch die Gesamtkommunikation und ihre Wirkung auf Wähler:innen erfasst werden. Konsument:innen / die Bevölkerung stellen ihre Ansichten, Gedanken und Erfahrungen auf diversen Sozialen Medien dar.

---

[27] Stieglitz/Dang-Xuan (2013); Hong/Nadler (2011).

[28] Hu/Liu (2012).

[29] Volkens et al. (2020).

Vorteile Sozialer Medien

Die Verbreitung und massenhafte Nutzung Sozialer Medien begann Mitte der 2000er-Jahre. Die Nutzung der verschiedenen Sozialen Medien steigt seitdem stetig an. Während sie zu Beginn auf regionale Verknüpfungen und vor allem ein junges Publikum begrenzt war, sind mittlerweile alle Altersgruppen, Geschlechter und gesellschaftlichen Schichten auf den Sozialen Medien miteinander verbunden.[30] Die flächendeckende Verbreitung von Smartphones ermöglicht die orts- und zeitunabhängige Nutzung durch die Konsument:innen/Wähler:innen.[31] Die Verbreitung der verschiedenen Medien unterscheidet sich je nach dem betrachteten Land zum Teil deutlich. Während in den Vereinigten Staaten der Großteil der Bevölkerung mindestens eine (teilweise) öffentliche Plattform nutzt, ist etwa in Deutschland die Nutzung nach wie vor in den jüngeren Altersgruppen weiter verbreitet. Zudem unterscheidet sich die Verbreitung der verschiedenen Plattformen. In Deutschland ist Facebook die am stärksten genutzte Plattform, auf den weiteren Rängen finden sich Twitter und stetig wachsend Instagram.[32] In den Vereinigten Staaten hingegen ist Twitter die meistgenutzte Plattform.[33] Ein weiterer Unterschied im internationalen Vergleich zeigt sich in der Nutzung der Plattformen. Während in den USA schnell eine kommerzielle und politische Nutzung vorangeschritten ist, ist in vielen europäischen Staaten die kommerzialisierte Nutzung der Plattformen ein vergleichsweise neues Phänomen.

Soziale Medien, als relativ neues Medium, stellen neue Herausforderungen an Auswertungsmethoden. Sie können jedoch durchaus in Zeitreihen angewandt werden.[34] Im politischen Kontext ist eine Auswertung seit Beginn der 2010er-Jahre möglich. In den USA werden seit Längerem über Mikrotargeting Wähler:innen gesucht, welche für Politiker:innen ansprechbar sind. In Europa sind derartige Auswertungen nicht mit den Datenschutzbestimmungen vereinbar. Die Anwendung von Mikrotargeting ist bereits länger möglich, da es sich um Einzelpersonen als Untersuchungsobjekte handelt.[35] Eine systematische Analyse zum Beispiel der politischen Kommunikation durch Entscheidungsträger:innen hingegen, bedarf der Aktivität von Parteien und Politiker:innen auf

30 Java et al. (2007).

31 Beisch/Schäfer (2020).

32 Ebenda.

33 Hu/Liu (2012).

34 Hu/Liu (2012).

35 Zuiderveen Borgesius et al. (2018); Papakyriakopoulos et al. (2017).

sozialen Kanälen. Die Notwendigkeit wurde von Parteien und Politiker:innen erst stückweise erkannt. Im deutschen Kontext ist die Wahlperiode 2013 bis 2017 die erste flächendeckend auswertbare Periode. Für die Bundestagswahl 2013 ist die kurzfristige Wahlkampfstrategie bereits analysierbar. Für Konsument:innen gelten ähnlich strenge Datenschutzbestimmungen zur Anonymisierung von Einzelpersonen. Die Datennotwendigkeiten für die Analyse von Konsumentscheidungen entsprechen jenen für politisches Mikrotargeting

Die Vorteile der Sozialen Medien liegen auf der Hand. Auf den Sozialen Medien kommunizieren Einzelpersonen und Unternehmen oder Politiker:innen ohne Zeitverzug. Abgesehen von redaktionell verfassten Beiträgen spiegeln die Beiträge die aktuellen Gedanken oder Bedürfnisse ohne Interpretationsschritte oder kognitive Verzerrung wider. Zudem können Konsument:innen/ Produzent:innen oder Wähler:innen/Parteien direkt miteinander in Kontakt treten und auf die jeweils andere Seite reagieren. Die stetige Produktion neuer Daten erlaubt eine tagesaktuelle Auswertung von Informationen. Zeitliche Restriktionen und Schlaglichter zu einem einzelnen Zeitpunkt können damit vermieden werden.

Einerseits haben die Sozialen Medien klare Vorteile, doch andererseits gibt es einige Faktoren, die die Analyse behindern. Zum einen sind die Beiträge zum Teil sehr kurz gehalten, zum Beispiel auf Twitter. Zum anderen sind häufig viele verschiedene Themen innerhalb eines Posts zusammengefasst. Eine klare Trennung beziehungsweise Zuordnung der Anteile eines Posts stellt sich teilweise als schwierig dar. Daneben erschwert die Formulierung der Nachrichten die Analyse, die Problematik ergibt sich aus der knappen und schnell verfassten Natur der Beiträge. Dabei werden unterschiedliche Abkürzungen für Wörter verwendet, Umgangssprache genutzt und es treten häufiger Rechtschreibungsfehler auf. Diese Punkte erschweren die systematische Auswertung und erfordern ein ausgeprägtes *pre-processing* der Rohdaten.[36] Abgesehen von offiziellen Accounts besteht die Möglichkeit, dass Personen sich auf Sozialen Medien nicht mit ihrer wahren Identität oder ihrem wahren Standort anmelden. Problematisch sind dabei insbesondere Bots, welche eine große Anzahl an Beiträgen automatisiert verfassen. Um diese Accounts zu filtern und nicht fälschlicherweise in ungewolltem Kontext in Analysen zu integrieren, lassen sich Algorithmen

[36] Hu/Liu (2012).

einsetzen. Bei einem Bewusstsein der Problematik gegenüber lässt sich somit eine fälschliche Integration verhindern.[37]

Anwendungen Sozialer Medien

Die Anwendungen zur textbasierten Auswertung sozialer Medien sind sehr vielschichtig. Die Möglichkeiten und Anwendungen erwachsen aus der hohen Nutzerzahl, der globalen Verfügbarkeit und dem vielfältigen Nutzerkreis. Während für die Analyse und Vorhersage der Börsen/Aktienkurse unter anderem öffentliche Profile und Statements der Zentralbanken und/oder Politiker:innen genutzt werden,[38] stehen bei „social emotion"- und „opinion mining"-Studien private Statements der gesamten Nutzerbasis im Mittelpunkt.[39] Aus der Betrachtung sozialer Emotionen lassen sich über automatisierte Verfahren zusätzlich Sentiment-Lexika entwickeln.[40] Darüber hinaus bieten sich die Sozialen Medien für die Erfassung von Events und Wendepunkten an.[41] Die Betrachtung des Konsument:innenverhaltens (siehe dazu nachfolgend den Beitrag in diesem Band) nutzt zumeist Sentiments und die gesamte Nutzerbasis eines definierten Georaums.[42] Ähnlich wie in diesem Beitrag werden die Sozialen Medien für die Analyse politischer Themen und des Parteiwettbewerbs genutzt.[43]

Eine konkrete Anwendung im Bereich der politischen Kommunikation analysiert das Kommunikationsverhalten deutscher Parteien in der Legislaturperiode 2013 bis 2017. Es wurden die offiziellen Facebook-Seiten der Bundesparteien sowie der jeweiligen Parteivorsitzenden und des Generalsekretärs genutzt (siehe Tabelle 1). Im Zeitraum von Januar 2014 bis Dezember 2017 konnte so die gesamte Wahlperiode inklusive der Koalitionsverhandlungen abgebildet werden. Für die gewählte Periode konnten zwischen 30 und 40 Themen pro Parteien identifiziert werden. Die Auswertung der Facebook-Beiträge erfolgte über

[37] Zum Beispiel Cai/Li/Zengi (2017).

[38] Nguyen/Shirai (2015); Besimi et al. (2019); Bollen/Mao/Zeng (2011); Oliveira/Cortez/Areal (2017).

[39] Rao et al. (2014); Bao et al. (2009); Vamshi/Pandey/Siva (2018).

[40] Deng et al. (2019); Xie/Li (2012).

[41] Xue et al. (2020); Qian et al. (2016).

[42] Homburg/Ehm/Artz (2015); Daas/Puts (2014); Pekar/Binner (2017).

[43] Joshi/Bhattacharyya/Carman; Takikawa/Nagayoshi (2017); Oliveira et al. (2018); Antonakaki et al. (2017).

Latent Dirichlet Allocation[44], eine unüberwachte Soft-Clustering-Methode. Die Daten wurden auf monatlicher Basis akkumuliert, um eine Aussage über das Kommunikationsverhalten, die Schwerpunkte und Veränderungen zu treffen. Die Parteienkommunikation wird separat ausgewertet, mit dem Ziel kleinere Themen oder Themendifferenzierungen der Parteien zu generieren.

Tab. 1: Parteien und Politker:innen im Datensatz

PARTEI		
AfD	Frauke Petry	Jörg Meuthen
CSU	Horst Seehofer	Andreas Scheuer
CDU	Angela Merkel	Peter Tauber
FDP	Christian Lindner	Nicola Beer
SPD	Sigmar Gabriel	Hubertus Heil
BÜNDNIS 90 / DIE GRÜNEN	Cem Özdemir	Simone Peter
DIE LINKE	Bernd Riexinger	Katja Kipping

Der Algorithmus gibt die prozentualen Anteile der einzelnen Themen innerhalb eines Beitrags und die wichtigsten Wörter eines Clusters aus. Zusätzlich wurde der Sentiment Score erstellt. Der Sentiment Score wird auf einer Skala von *–1* bis *+1*, auf der Grundlage eines Wörterbuchs errechnet. Für das Wörterbuch wurden von Experten Wörter hinsichtlich ihrer positiven oder negativen Konnotation kodiert.[45] Der Sentiment Score für ein Dokument i mit den Sentiment-Wortwerten w und dem Gesamtvokabular v errechnet sich wie in der folgenden Formel dargestellt.

$$S_i = \frac{\sum w_{pos,i} - \sum w_{neg,i}}{\sum v_i}$$

Im Ergebnis konnten signifikante und den Parteien entsprechende Ergebnisse erzielt werden. Das Profil der Parteien, insbesondere der kleinen Parteien, entspricht den Schwerpunktsetzungen in den Wahlprogrammen. Der Vorteil anderen Ansätzen gegenüber ist die monatliche Auswertbarkeit und darauf aufbauend die Möglichkeit von Analysen der Reaktionen von Wähler:innen. Beispielhaft werden an dieser Stelle die Ergebnisse zur Arbeitsmarktpolitik, zur Sozial- und Familienpolitik, zur wirtschaftlichen Entwicklung Deutschlands sowie zu außergewöhnlichen Ereignissen und deren Erkennung betrachtet.

---

[44] Blei/Ng/Jordan (2003).

[45] Remus/Quasthoff/Heyer (2010).

Die Anwendung von Latent Dirichlet Allocation (LDA) gibt als Resultat zwei Ergebnisse aus. Zum einen werden die Themencluster in Form der häufigsten Wörter ausgegeben. Zum anderen wird für jeden Text der zugehörige prozentuale Anteil an einem Thema ausgegeben. Durch die Definition der Parameter kann festgelegt werden, ob viele oder wenige Themen pro Nachricht erwartet werden. Im nachfolgenden Fall wurde, ausgehend von dem tendenziell kurzen Charakter der Facebook-Posts angenommen, dass wenige Themen pro Nachricht kommuniziert werden. Für die Interpretation ist relevant, dass ein Thema, welches in den Clustern nicht auftaucht, nicht zwangsläufig gar nicht behandelt wird. Es wird jedoch in entweder deutlich untergeordneter Relevanz oder in nicht differenziertem Vokabular kommuniziert.

Tab. 2: Codebuch – Auszug zu Sozialpolitik und Budget, Wachstum und Entwicklung

SOZIALPOLITIK

Arbeitsmarkt	Arbeitslosigkeit Löhne Arbeitslosengeld (Hartz IV) Soziale Sicherungssysteme Arbeitsrecht Gewerkschaften Start-ups
Rentenpolitik	Rentenversicherung Altersarmut
Familienpolitik	Work-Life-Balance Mütterrente Elterngeld Kinderarmut
Pflege	
Wohnraum	
Bildungspolitik	

BUDGET, WACHSTUM, ENTWICKLUNG

Budget	Schwarze Null Investitionen
Vermögen	Vermögenssteuer
Wachstum	Wachstumsfaktoren Prosperity
Freihandel	TTIP CETA
Digitalisierung	
Ländliche Entwicklung	

Die resultierenden Wortlisten müssen kodiert werden. Auf Basis eines detaillierten Codebuchs können die Listen Themen zugeordnet werden. Beispielhaft sind in Tabelle 2 die Abschnitte des Codebuchs für die Themenfelder Sozialpolitik, Budget, Wachstum und Entwicklung abgebildet. Das Codebuch wurde in einem explorativen Ansatz entwickelt, basierend auf den Resultaten der LDA-Cluster. Der Vorteil der automatisierten Textanalyse liegt im frühzeitigen Erkennen neuer Themenfelder, daher ist ein explorativer Ansatz der Themengenerierung sinnvoll.

Tab. 3: Themenliste – Arbeitsmarkt

DIE LINKE			
ARBEITSRECHT	GEWERKSCHAFTEN	LÖHNE	ARBEITSLOSIGKEIT
„arbeit“	„beschaeftigt“	„euro“	„hartz“
„gut“	„gewerkschaft“	„mindestlohn“	„sanktion“
„leiharbeit“	„gut“	„million“	„betroff“
„beschaeftigt“	„arbeit“	„jahr“	„abschaff“
„mensch“	„verdi“	„milliard“	„grundrecht“
„leb“	„unterstuetz“	„unternehm“	„jobcent“
„befrist“	„streik“	„zahl“	„sanktionsfrei“
„job“	„loehn“	„ausnahm“	„mindestsicher“
„preka“	„amazon“	„niedrig“	„bundesregier“
"gleich"	„metall“	„fordert“	„andrea“

SPD	CSU	CDU	
LOHNE/ GEWERKSCHAFTEN	ARBEITSLOSIGKEIT/ LÖHNE	LÖHNE	ARBEITSLOSIGKEIT
„mindestlohn“	„loewenstark“	„sich“	„deutschland“
„rent“	„gut“	„stark“	„gut“
„arbeit“	„best“	„deutschland“	„mensch“
„mensch“	„digital“	„inn“	„wirtschaft“
„gut“	„wirtschaft“	„wirtschaft“	„jahr“
„jahr“	„prozent“	„arbeit“	„arbeitslos“
„gesetz“	„arbeitslosenquot“	„bleibt“	„zahl“
„prozent“	„arbeitsmarkt“	„wohlstand“	„arbeit“
„arbeitnehm“	„freistaat“	„arbeitsplaetz“	„prozent“
„andrea“	„top“	„zeit“	„arbeitsmarkt“

Die Parteien diskutieren in Bezug auf die wichtigen wirtschafts- und sozialpolitischen Themenfelder die gleichen Oberthemen. Die Unterscheidung wird in der Diskurssetzung, also den präziseren Unterthemen, evident. Dies wird bei sozialpolitischen Themen deutlich. Tabelle 2 weist die zehn Wörter aus, die im Rahmen von arbeitsmarktbezogenen Themen am häufigsten verwendet wurden. Die zuvor beschriebene Unterteilung der Themen wird hier verdeutlicht. Darüber hinaus zeigt das Agenda-Setting sowohl den Unterschied zwischen Mitte-Links- und Mitte-Rechts-Parteien als auch die Konzentration auf eine kontroverse versus eine ergebnisorientierte Debattenstruktur.

Während in der Linken ein breites Spektrum von Arbeitsmarktfragen unter besonderer Berücksichtigung der sozialen Gerechtigkeit diskutiert wird, konzentriert sich die CDU auf Erfolge und gute Arbeitsmarktbedingungen im Sinne einer niedrigen Arbeitslosigkeit. Die Unterschiede zwischen Mitte-Links- und Mitte-Rechts-Parteien werden auch im Vergleich von CDU und SPD sichtbar. Beide Parteien gingen 2013 eine Koalition ein, wobei sich die SPD auf zentrale sozialdemokratische Themen wie den Mindestlohn konzentriert, der jedoch von der CDU nicht erwähnt wird. Das obige Beispiel zeigt, zu welchen sinnvollen Ergebnissen die Anwendung von Themenmodellen auf die offizielle Social-Media-Kommunikation der Parteien führt. Auffällig ist bei den weiteren Themen der sozialpolitischen Kommunikation auch die Zuordenbarkeit der Themen zu Parteien oder mindestens die Regierungszugehörigkeit oder das politische Spektrum.

So werden zum Beispiel von allen Parteien, CDU und CSU ausgenommen, die Freihandelsabkommen TTIP und CETA behandelt. Während die Linke und die Grünen den Abkommen eher ablehnend gegenüberstehen, kann aus den zehn häufigsten Wörtern bei SPD und AfD keine klare Positionierung abgeleitet werden. Die FDP kommt in den zehn häufigsten Wörtern eher zu einem positiven Urteil. Linke und Grüne diskutieren zudem ihre Kernthemen Umverteilung respektive Umweltschutz im Zusammenhang mit der wirtschaftlichen Entwicklung Deutschlands. Bei den Regierungsparteien und der FDP kommt die Digitalisierung hinzu. Während bei SPD und CDU der Fortschritt im Tenor eher positiv klingt, deutet die Kommunikation der FDP auf kritische Aspekte in Bezug auf die Digitalisierungsstrategie hin. Insgesamt zeichnen die wirtschaftspolitischen Themen, ähnlich wie die sozialpolitischen Themen, das Profil der Parteien und ihrer Ausrichtung nach.

Tab. 4: Themenliste – Budget, Wachstum und Entwicklung

	DIE LINKE		AFD		
	FREIHANDEL	VERMÖGENSTEUER	TTIP	WACHSTUM	SCHWARZE NULL
1	„ttip“	„reich“	„buerg“	„deutschland“	„euro“
2	„ceta“	„deutschland“	„ttip“	„jahr“	„geld“
3	„stopp“	„prozent“	„bargeld“	„zahl“	„jahr“
4	„polit“	„gerecht“	„freiheit“	„prozent“	„kost“
5	„konzern“	„sozial“	„lehnt“	„studi“	„milliard“
6	„freihandelsabkomm“	„reichtum“	„oeffent“	„polit“	„million“
7	„demokrati“	„vermoeg“	„kontroll“	„mensch“	„steuerzahl“
8	„aktion“	„land“	„staat“	„deutsch“	„pro“
9	„mensch“	„einkomm“	„abschaff“	„articl“	„hoeh“
10	„abkomm“	„gesellschaft“	„gross“	„wirtschaft“	„steu“

	BÜNDNIS 90 / DIE GRÜNEN		SPD		
	WACHSTUMS-FAKTOREN	FREIHANDEL	WACHSTUMS-FAKTOREN	TTIP	DIGITALISIE-RUNG
1	„gerecht“	„ttip“	„wirtschaft“	„ttip“	„digital“
2	„sozial“	„ceta“	„unternehm“	„bundesregier“	„frag“
3	„oekolog“	„umwelt“	„deutschland“	„ceta“	„zukunft“
4	„gesellschaft“	„fair“	„deutsch“	„kommission“	„digitalleb“
5	„wirtschaft“	„stopp“	„industri“	„freihandelsab-komm“	„arbeit“
6	„weltoff“	„handel“	„jahr“	„bmwi“	„digitalisier“
7	„umwelt“	„handelsabkomm“	„arbeitsplaetz“	„entscheid“	„diskuti“
8	„zentral“	„usa“	„zukunft“	„oeffent“	„gesellschaft“
9	„arbeit“	„mensch“	„stark“	„debatt“	„leb“
10	„zusammenhalt“	„konzern“	„gut“	„verhandl“	„thema“

	CSU		CDU		
	LÄNDLICHE ENTWICKLUNG	INVESTITIONEN	SCHWARZE NULL	DIGITALISIE-RUNG	LÄNDLICHE ENTWICKLUNG
1	„gross“	„euro“	„jahr“	„digital“	„gefuehrt“
2	„polit“	„bund“	„schuld“	„digitalisier“	„bundesregier“
3	„zeit“	„jahr“	„haushalt“	„thema“	„euro“
4	„gemeinsam“	„laend“	„bundestag“	„diskutiert“	„gut“
5	„wichtig“	„milliard“	„schwarz“	„chanc“	„infos“
6	„herausforder“	„zukunft“	„forschung“	„wirtschaft“	„milliard“
7	„bleib“	„schuld“	„bildung“	„frag“	„pfleg“
8	„gut“	„bundesla“	„solid“	„wandel“	„nachricht“
9	„dialog“	„finanz“	„null“	„mittelpunkt“	„internet“
10	„gespraech“	„entlast“	„betont“	„diskuti“	„bund“

	FDP			
	BUDGET	TTIP	DIGITALISIERUNG	START-UPS
1	„steu“	„wirtschaft“	„deutschland“	„bess“
2	„entlast“	„deutsch“	„digitalisier“	„deutschland“
3	„soli“	„ttip“	„digital“	„arbeit“
4	„wissing“	„arbeitsplaetz“	„chanc“	„gruend“
5	„euro“	„gut“	„bildung“	„ide“
6	„politikdierech-nenkann“	„gerad“	„braucht“	„unternehm“
7	„kalt“	„unternehm“	„endlich“	„buerokrati“
8	„progression“	„handwerk“	„brauch“	„fordert“
9	„buerg“	„mindestlohn“	„infrastruktur“	„gerad“
10	„geld“	„wohlstand“	„agenda“	„braucht“

Nachdem die Themencluster codiert sind, kann der prozentuale Anteil der Kommunikation an den Themenclustern ermittelt werden. Um eine aussagekräftige Verteilung zu erzielen, werden hierfür die Oberthemen Sozialpolitik und Budget, Wachstum und Entwicklung in der Darstellung kumuliert. Für eine präzisere Auswertung und konkrete Anwendungen im Bereich der Wählerreaktion können die detaillierteren Themenfelder wie Arbeitsmarkt oder Familienpolitik genutzt werden.

In Abbildung 2 wird die kumulierte Kommunikation der Bundesparteien zu sozialpolitischen Themen (zur Ausrichtung der Debatte siehe Tabelle 4) dargestellt. Exemplarisch zeigen die sozialpolitischen Themen, dass die Parteien bezüglich der kommunizierten Themen starke Schwerpunkte setzen. Über den gesamten Zeitraum hinweg bestreitet die Linke das Thema in ausgeprägter Form. Die regierungsbeteiligten Parteien SPD und CDU folgen mit einer gleichbleibend hohen Beteiligung am Themenspektrum, verweisen dabei jedoch hauptsächlich auf arbeitsmarktbezogene Themen und im Fall der CDU auf Erfolge der Regierung.

Am rechten Rand des politischen Spektrums wird der Akzent auf migrationsbezogene Themen gesetzt und als Wahlkampfthema im September 2017 weiter intensiviert. Abbildung 3 zeigt zwei Phänomene deutlich auf. Zum einen ist ein Anstieg der Kommunikation, ausgelöst durch ein externes Ereignis, zu sehen (Anstieg der Zuwanderung Geflüchteter 2015). Zum anderen kann, ähnlich wie in der sozialpolitischen Kommunikation, die Vereinnahmung eines Themas durch einzelne Parteien beobachtet werden: Die Kommunikation der CSU und der AfD über zuwanderungsrelevanten Themen nimmt zu.

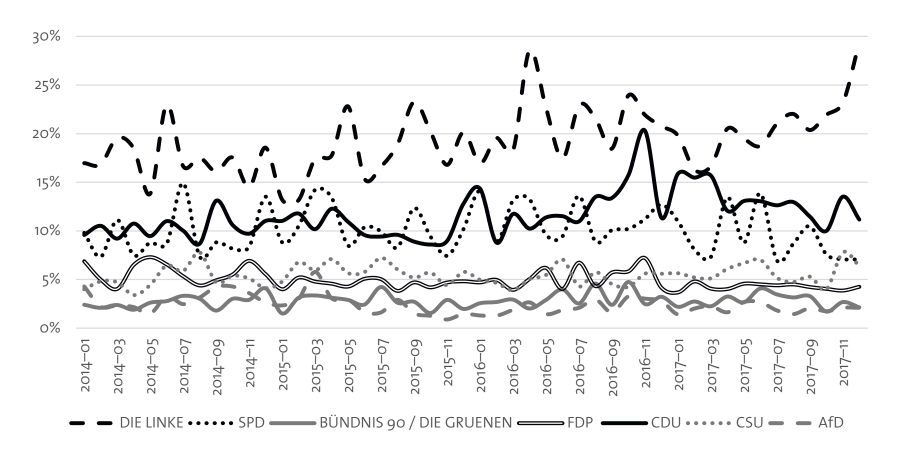


Abb. 2: Monatlich kumulierte Kommunikation – Sozialpolitik
Quelle: eigene Darstellung.

Ebenfalls zu beobachten ist in der sozialpolitischen Kommunikation, dass die Vereinnahmung eines Themas mit einer problemorientierten Darstellung der Thematik einhergeht. Während insbesondere das linke Parteienspektrum die Integration und die Fluchtursachen thematisiert, werden von CSU und AfD verstärkt die Konsequenzen und problematischen Aspekte der Zuwanderung dargestellt. Des Weiteren ist bei der AfD zu beobachten, dass, ähnlich wie bei der Partei DIE LINKE, das Thema in vielen Varianten kommuniziert wird. Als interessant zu vermerken ist zudem, dass der Wechsel an der AfD-Parteispitze (5. Juli 2015) und damit die Neuausrichtung deutlich zu erkennen ist, Euro-kritische Themen werden zu migrationskritischen Themen. Dies ist umso bezeichnender, als über den gesamten Zeitraum dieselben Personen in die Analyse einbezogen waren.

Die Analyse der Relevanz einzelner Themen zeigt drei zu erwartende und wichtige Phänomene der politischen Kommunikation. Erstens: Kleine und Oppositionsparteien verfolgen die Strategie, einzelne Themen für sich zu vereinnahmen und den Diskurs zu bestimmen.[46] Zweitens können externe Effekte mit politischer Durchschlagskraft in der öffentlichen Kommunikation zuverlässig erkannt werden. Drittens werden Strategiewechsel, innerparteiliche Auseinandersetzungen und wahlkampfstrategische Überlegungen evident.

---

[46] Die Ergebnisse werden von verschiedenen Studien zur strategischen Ausrichtung von Parteien unterstützt (siehe zum Beispiel Dragu/Fan [2016]; Glazer/Lohmann [1999]).

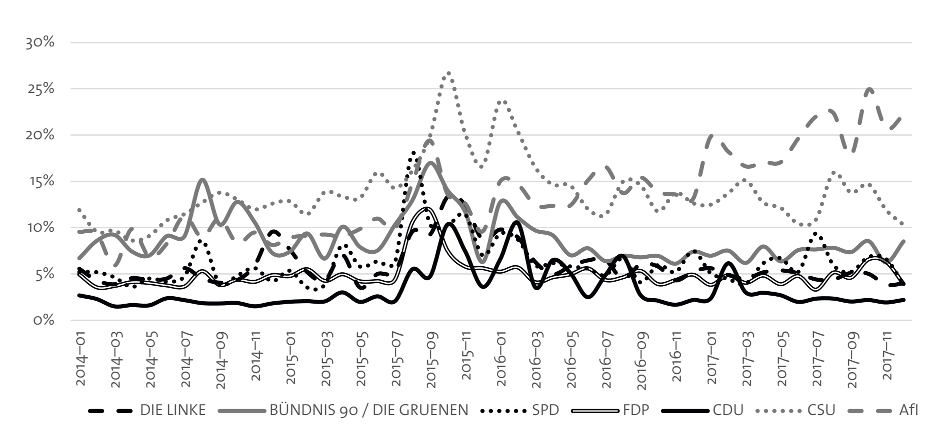


Abb. 3: Monatlich kumulierte Kommunikation – Migration
Quelle: eigene Darstellung.

Um eine empirische Auswertung der Daten zu ermöglichen, muss der Tenor der Parteienkommunikation quantifiziert werden. Verschiedene Ansätze können hierfür verwandt werden. Die einfachste Möglichkeit ist die Verwendung eines wörterbuchbasierten Verfahrens zur Berechnung der Sentiments einzelner Wörter. Ein Problem dieser Methodik liegt in der Ausgestaltung des Wörterbuchs. Gerade Wortneuschöpfungen mit einem stark positiven oder negativen Charakter sind in den Wörterbüchern selten enthalten. Einen Eindruck von der Kommunikation spiegelt die Methode jedoch gut wider. Exemplarisch sind an dieser Stelle die Oppositionsparteien (Die Linke, Bündnis 90 / die Grünen, FDP, AfD) sowie die Regierungsparteien dargestellt. Oppositionsparteien fallen durch ihre weitestgehend negativen Sentiment-Werte auf (Abbildung 4). Die Regierungsparteien hingegen sind überwiegend im positiven Wertebereich zu finden. Dies deutet auf eine sehr unterschiedliche Darstellung der Lage in verschiedenen Themenbereichen und die Hervorhebung der Erfolge auf Seiten der Regierungsparteien hin.

Weitere Ansätze zur Analyse beinhalten maschinelle Ansätze zum NLP oder psychologische Ansätze zur Generierung eines psychologischen Motivs. Diese Methodik wird im letzten Beitrag in diesem Band genauer betrachtet.

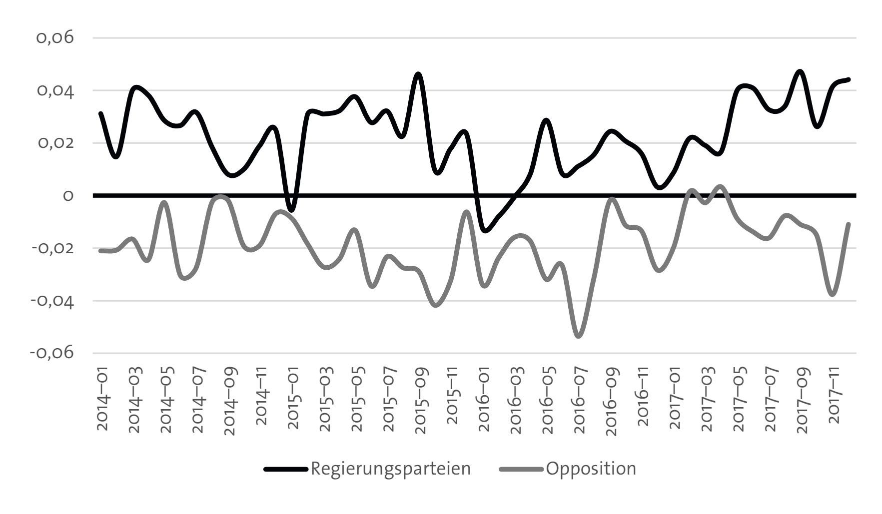


Abb. 4: Kumulierte Sentiment Scores – Regierungs- und Oppositionsparteien
Quelle: eigene Darstellung.

Die Ergebnisse zeigen die vielfältigen Möglichkeiten auf, mit denen textbasierte Verfahren in der politischen Analyse eingesetzt werden können. Es können frühzeitig relevante Themen erkannt und ihre Wirkung auf Wähler:innen und Konsument:innen untersucht werden. Die Methodik erlaubt in weitestgehend automatisierter Form, die Profile von Parteien zu analysieren. Dies lässt sich auf vielfältige weitere Themenfelder übertragen. Neben politischer Kommunikation und wirtschaftspolitischen Aspekten können Trends der gesamtwirtschaftlichen Entwicklung ausgewertet werden. Textdaten sollten daher in Forschung und Wirtschaft einen höheren Stellenwert erhalten. Sie können quantitative Ergebnisse unterstützen beziehungsweise sogar deren Aussagegehalt erhöhen.

Schlussfolgerungen

Wie gezeigt wurde, erfordert die Verwendung unstrukturierter Textdaten komplexe Auswertungsmethoden. Die Methoden können sich je nach Zielstellung unterscheiden. Gemeinsam ist ihnen allen das Ziel, mit möglichst geringen händischen Kodierungen große Textmengen zu analysieren. Während bei überwachten Strategien die Kodierung eines Trainingsdatensets nötig ist, ist bei un-

überwachten Strategien eine nachträgliche Kodierung der Cluster notwendig. Wie im Abschnitt „Texte als Daten“ gezeigt, haben beide Methoden Vorteile.

Gerade für den in diesem Beitrag erläuterten Ansatz zur Analyse politischer Kommunikation bietet sich das unüberwachte Clusteringverfahren an. Die explorative Strategie kann kann dazu dienen, neue Themen zu erkennen und so gesellschaftliche Entwicklungen für spätere Analyseverfahren aufzudecken. Das ermöglicht es, in Fore- und Nowcasts Entwicklungen abzubilden,[47] welche in Surveys noch nicht als implementierte Frage integriert sind. Neben Clusteringmethoden können über Sentimentanalysen oder präzisere psychologische Tools Informationen zu Stimmungen extrahiert werden.

Verschiedene Veröffentlichungen zeigen bereits den Nutzen unstrukturierter Daten in empirischen Verfahren. So konnten Entwicklungen an Aktienmärkten, die Reaktion auf Zentralbankkommunikation und positive Zusammenhänge zwischen Konsument:innenverhalten und Kommunikation in den Sozialen Medien identifiziert werden.

Die Ergebnisse und die vielfältigen Anwendungsfelder zeigen das Potenzial unstrukturierter Daten auf. Die Entwicklung zeigt in Richtung eines sich stetig ausweitenden Datenpools, eine Auswertung scheint daher ein vielversprechender und notwendiger Weg zu sein. Da die Methoden noch nicht gleichwertig validiert sind, bleibt das Feld ausgesprochen dynamisch, und es ist mit zunehmend besseren Modellen und Ergebnissen zu rechnen.

Literatur

Aggarwal, Charu C. (2012): An Introduction to Text Mining. in: Mining Text Data. hrsg. von Aggarwal, Charu C. / Zhai, ChengXiang, New York, NY 2012, S. 1–10.

Aggarwal, Charu C. / Zhai, ChengXiang (2012a): A Survey of Text Classification Algorithms. in: Mining Text Data. hrsg. von Aggarwal, Charu C. / Zhai, ChengXiang, New York, NY 2012, S. 163–222.

Aggarwal, Charu C. / Zhai, ChengXiang (2012b): A Survey of Text Clustering Algorithms. in: Mining Text Data. hrsg. von Aggarwal, Charu C. / Zhai, ChengXiang, New York, NY 2012, S. 77–128.

---

[47] Für weitere Informationen zu Fore- und Nowcasts siehe den vorhergehenden Beitrag in diesem Band.

Antonakaki, Despoina / Spiliotopoulos, Dimitris / V Samaras, Christos / Pratikakis, Polyvios / Ioannidis, Sotiris / Fragopoulou, Paraskevi (2017): Social media analysis during political turbulence. in: PloS one, Vol. 12, Nr. 10 (2017), e0186836.

Bao, Shenghua / Xu, Shengliang / Zhang, Li / Yan, Rong / Su, Zhong / Han, Dingyi / Yu, Yong (2009): Joint Emotion-Topic Modeling for Social Affective Text Mining. in: Data Mining, 2009. ICDM, 2009 Ninth IEEE International Conference, S. 699–704.

Beisch, Natalie / Schäfer, Carmen (2020): Internetnutzung mit großer Dynamik: Medien, Kommunikation, Social Media. in: Media Perspektiven, Nr. 9 (2020), S. 462–481.

Besimi, Adrian / Dika, Zamir / Shehu, Visar / Selimi, Mubarek (2019): Applied Text-Mining Algorithms for Stock Price Prediction Based on Financial News Articles. in: Managing Global Transitions, Vol. 17, Nr. 4 (2019), S. 335–351.

Bishop, Christopher M. (2006): Pattern recognition and machine learning. 1. Aufl., New York 2006.

Blei, David M. / Lafferty, John D. (2006): Dynamic topic models. in: ICML 2006. Proceedings, twenty-third International Conference on Machine Learning. hrsg. von Moore, Andrew / Cohen, William W., New York 2006, S. 113–120.

Blei, David M. / Lafferty, John D. (2007): A correlated topic model of Science. in: The annals of applied statistics Vol. 1, Nr. 1 (2007), S. 17–35.

Blei, David M. / Lafferty, John D. (2009): Topic Models. in: Text mining. Classification, clustering, and applications. hrsg. von Srivastava, Ashok N. / Sahami, Mehran, Boca Raton, Fla., 2009, S. 71–93.

Blei, David M. / Ng, Andrew Y. / Jordan, Michal I. (2003): Latent dirichlet allocation. in: Journal of Machine Learning Research, Vol. 3, Nr. 0 (2003), S. 993–1022.

Bollen, Johan / Mao, Huina / Zeng, Xiaojun (2011): Twitter mood predicts the stock market. in: Journal of Computational Science, Vol. 2, Nr. 1 (2011), S. 1–8.

Cai, Chiyu / Li, Linjing / Zengi, Daniel (2017): Behavior enhanced deep bot detection in social media. in: 2017 IEEE International Conference on Intelligence and Security Informatics (ISI), S. 128–130.

Conference Board: Leading Economic Index. elektronisch veröffentlicht unter der URL: https://conference-board.org/data/bcicountry.cfm?cid=4, 02.11.2020.

Daas, Piet J. H. / Puts, Marco J. H. (2014): Social media sentiment and consumer confidence, Frankfurt, 2014.

Deng, Dong / Jing, Liping / Yu, Jian / Sun, Shaolong / Ng, Michael K. (2019): Sentiment lexicon construction with hierarchical supervision topic model. in: IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 27, Nr. 4 (2019), S. 704–718.

Dragu, Tiberiu / Fan, Xiaochen (2016): An Agenda-Setting Theory of Electoral Competition. in: The Journal of Politics, Vol. 78, Nr. 4 (2016), S. 1170–1183.

Einav, Liran / Levin, Jonathan (2014): Economics in the age of big data. in: Science, Vol. 346, Nr. 6210 (2014), S. 715–723.

Gentzkow, Matthew / Kelly, Bryan / Taddy, Matt (2019): Text as Data. in: Journal of Economic Literature, Vol. 57, Nr. 3 (2019), S. 535–574.

GfK: GfK-Konsumklima MAXX. elektronisch veröffentlicht unter der URL: https://www.gfk.com, 02.11.2020.

Glazer, Amihai / Lohmann, Susanne (1999): Setting the Agenda: Electoral Competition, Commitment of Policy, and Issue Salience. in: Public Choice, Vol. 99, Nr. 3/4 (1999), S. 377–394.

Hofmann, Thomas (2013): Probabilistic Latent Semantic Analysis, in: Machine Learning, Vol. 42, Nr. 1–2, S. 177–196.

Homburg, Christian / Ehm, Laura / Artz, Martin (2015): Measuring and Managing Consumer Sentiment in an Online Community Environment. in: Journal of Marketing Research, Vol. 52, Nr. 5 (2015), S. 629–641.

Hong, Sounman / Nadler, Daniel (2011): Does the early bird move the polls? in: dg.o 2011 : the proceedings of the 12th annual International Digital Government Research Conference : Digital Government Innovation in Challenging Times : University of Maryland, College Park, Maryland, USA, June 12–15, 2011. hrsg. von Luna Reyes, Luis F. / Chun, Soon A. / Bertot, John, 2011, S. 182.

Hu, Xia / Liu, Huan (2012): Text Analytics in Social Media. in: Mining Text Data. hrsg. von Aggarwal, Charu C. / Zhai, ChengXiang, New York, NY 2012, S. 385–414.

Iyetomi, Hiroshi / Aoyama, Hideaki / Fujiwara, Yoshi / Souma, Wataru / Vodenska, Irena / Yoshikawa, Hiroshi (2020): Relationship between Macroeconomic Indicators and Economic Cycles in U.S. in: Scientific reports, Vol. 10, Nr. 8420 (2020), S. 1–12.

Java, Akshay / Song, Xiaodan / Finin, Tim / Tseng, Belle (2007): Why we twitter. in: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis – WebKDD/SNA-KDD '07. hrsg. von Zhang, Haizheng / Mobasher, Bamshad / Giles, Lee / McCallum, Andrew / Nasraoui, Olfa / Spiliopoulou, Myra / Srivastava, Jaideep / Yen, John, New York, New York, USA 2007, S. 56–65.

Joshi, Aditya / Bhattacharyya, Pushpak / Carman, Mark (2016): Political Issue Extraction Model: A Novel Hierarchical Topic Model That Uses Tweets By Political And Non-Political Authors. in: Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Stroudsburg, PA, USA, S. 82–90.

Kumar, P. / Das, T. K. (2013): BIG Data Analytics: A Framework for Unstructured Data Analysis. in: International journal of engineering and technology, Vol. 5 (2013), S. 153–156.

Liu, Bing / Zhang, Lei (2012): A Survey of Opinion Mining and Sentiment Analysis. in: Mining Text Data. hrsg. von Aggarwal, Charu C. / Zhai, ChengXiang, New York, NY 2012, S. 415–463.

Murphy, Kevin P. (2012): Machine learning. A probabilistic perspective, Cambridge, MA 2012.

Nenkova, Ani / McKeown, Kathleen (2012): A Survey of Text Summarization Techniques. in: Mining Text Data. hrsg. von Aggarwal, Charu C. / Zhai, ChengXiang, New York, NY 2012, 43–76.

Nguyen, Thien H. / Shirai, Kiyoaki (2015): Topic modeling based sentiment analysis on social media for stock market prediction. in: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. hrsg. von Zong, Chengqing / Strube, Michael, 2015, S. 1354–1364.

OECD (2017): Leading indicators. Consumer Confidence Index 2017.

Oliveira, Lucas / Vaz de Melo, Pedro / Amaral, Marcelo / Pinho, José Antônio (2018): When Politicians Talk About Politics: Identifying Political Tweets of Brazilian Congressmen. in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 12, Nr. 1 (2018), S. 664–667.

Oliveira, Nuno / Cortez, Paulo / Areal, Nelson (2017): The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. in: Expert Systems with Applications, Vol. 73 (2017), S. 125–144.

Papakyriakopoulos, Orestis / Shahrezaye, Morteza / Thieltges, Andree / Medina Serrano, Juan Carlos / Hegelich, Simon (2017): Social Media und Microtargeting in Deutschland. in: Informatik-Spektrum, Vol. 40, Nr. 4 (2017), S. 327–335.

Pekar, Viktor / Binner, Jane (2017): Forecasting Consumer Spending from Purchase Intentions Expressed on Social Media. in: Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Stroudsburg, PA, USA 2017, S. 92–101.

Qian, Shengsheng / Zhang, Tianzhu / Xu, Changsheng / Shao, Jie (2016): Multi-Modal Event Topic Model for Social Event Analysis. in: IEEE Transactions on Multimedia, Vol. 18, Nr. 2 (2016), S. 233–246.

Quinn, Kevin M. / Monroe, Burt L. / Colaresi, Michael / Crespin, Michael H. / Radev, Dragomir R. (2010): How to Analyze Political Attention with Minimal Assumptions and Costs. in: American Journal of Political Science, Vol. 54, Nr. 1 (2010), S. 209–228.

Rao, Yanghui / Li, Qing / Wenyin, Liu / Wu, Qingyuan / Quan, Xiaojun (2014): Affective topic model for social emotion detection. in: Neural Networks, Vol. 58 (2014), S. 29–37.

Ravi, Kumar / Ravi, Vadlamani (2015): A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. in: Knowledge-Based Systems, Vol. 89 (2015), S. 14–46.

Remus, Robert / Quasthoff, Uwe / Heyer, Gerhard (2010): SentiWS – A Publicly Available German-language Resource for Sentiment Analysis. in: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta 2010.

Roberts, Margaret E. / Stewart, Brandon M. / Tingley, Dustin / Lucas, Christopher / Leder-Luis, Jetson / Gadarian, Shana K. / Albertson, Bethany / Rand, David G. (2014): Structural topic models for open-ended survey responses. in: American Journal of Political Science, Vol. 58, Nr. 4 (2014), S. 1064–1082.

Sakaki, Takeshi / Okazaki, Makoto / Matsuo, Yutaka (2010): Earthquake shakes Twitter users. in: Proceedings of the 19th international conference on World wide web – WWW '10, New York, New York, USA 2010, S. 851.

Stieglitz, Stefan / Dang-Xuan, Linh (2013): Social media and political communication: a social media analytics framework. in: Social network analysis and mining, Vol. 3, Nr. 4 (2013), S. 1277–1291.

Takikawa, Hiroki / Nagayoshi, Kikuko (2017): Political polarization in social media: Analysis of the „Twitter political field" in Japan. in: 2017 IEEE International Conference on Big Data (Big Data), Piscataway 2017, S. 3143–3150.

Tetlock, Paul C. (2007): Giving Content to Investor Sentiment: The Role of Media in the Stock Market. in: The Journal of Finance, Vol. 62, Nr. 3 (2007), S. 1139–1168.

Vamshi, Krishna B. / Pandey, Ajeet Kumar / Siva, Kumar A. P. (2018): Topic Model Based Opinion Mining and Sentiment Analysis. in: 2018 International Conference on Computer Communication and Informatics. January 04-06, 2018, Coimbatore, India. hrsg. von Informatics, I. C. o. C. C. a., Piscataway, NJ, 2018, S. 1–4.

Volkens, Andrea / Burst, Tobias / Krause, Werner / Lehmann, Pola / Matthieß, Theres / Merz, Nicolas / Regel, Sven / Weßels, Bernhard / Zehnter, Lisa / Wissenschaftszentrum Berlin Für Sozialforschung (WZB; 2020): Manifesto Project Dataset.

WorldBank (2019): Individuals using the internet (% of population). elektronisch veröffentlicht unter der URL: https://data.worldbank.org/indicator/IT.NET.USER.ZS, 5.1.2021.

Xie, Rui / Li, Chunping (2012): Lexicon construction: A topic model approach. in: 2012 International Conference on Systems and Informatics. ICSAI 2012: Yantai, Shandong, China, 19–20 May 2012, Piscataway, NJ 2012, S. 2299–2303.

Xue, Feng / Hong, Richang / He, Xiangnan / Wang, Jianwei / Qian, Shengsheng / Xu, Changsheng (2020): Knowledge-Based Topic Model for Multi-Modal Social Event Analysis. in: IEEE Transactions on Multimedia, Vol. 22, Nr. 8 (2020), S. 2098–2110.

Zuiderveen Borgesius, Frederik / Möller, Judith / Kruikemeier, Sanne / Ó Fathaigh, Ronan / Irion, Kristina / Dobber, Tom / Bodo, Balazs / Vreese, Claes H. de (2018): Online political microtargeting: promises and threats for democracy. in: Utrecht Law Review, Vol. 14, Nr. 1 (2018), S. 82–96.

Onlinedaten und Konsumentscheidungen

Voraussagen anhand von Daten aus Social Media und Suchmaschinen

Deniz Dilan Karaman Örsal

Einleitung

Dieser Beitrag beschäftigt sich mit dem Einsatz von Onlinedaten bei der Prognose grundlegender makroökonomischer Indikatoren wie privatem Konsum und Konsumentenvertrauensindikatoren. Dabei werden aktuelle Veröffentlichungen zusammengefasst und neue Anwendungsmöglichkeiten von Onlinedaten aufgezeigt.

Mit der breiten Nutzung des Internets und der rasanten Entwicklung der Digitalisierung stehen in den letzten Jahren neue und große Datenquellen zur Verfügung. Mit Hilfe von Kurznachrichten, überwiegend aus den sozialen Netzwerken (wie zum Beispiel Twitter, Facebook, Instagram, YouTube), oder aggregierten Indizes, die anhand von Suchanfragen konstruiert werden (wie unter anderem Google Trends, Google Insights), werden neue Vorhersagemodelle entwickelt, um wirtschaftliche Indikatoren frühzeitig beziehungsweise in Echtzeit zu prognostizieren.

In der Literatur wird auf einen Zusammenhang zwischen der Stimmung der Konsument:innen und dem privaten Konsum hingewiesen. Da der private Konsum 60 % bis 70 % des Bruttoinlandsprodukts ausmacht, können Veränderungen in der Verbraucherstimmung Veränderungen im privaten Konsum signalisieren.

Die Stimmung der Verbraucher:innen wird im Verlauf von Rezessionen akribisch beobachtet, da jede wesentliche Änderung oder jedes Fehlen einer Änderung als wichtiges Zeichen eines nahe gelegenen Wendepunkts oder einer Verlängerung des Tiefs angesehen wird.[1]

[1] Vuchelen (2004).

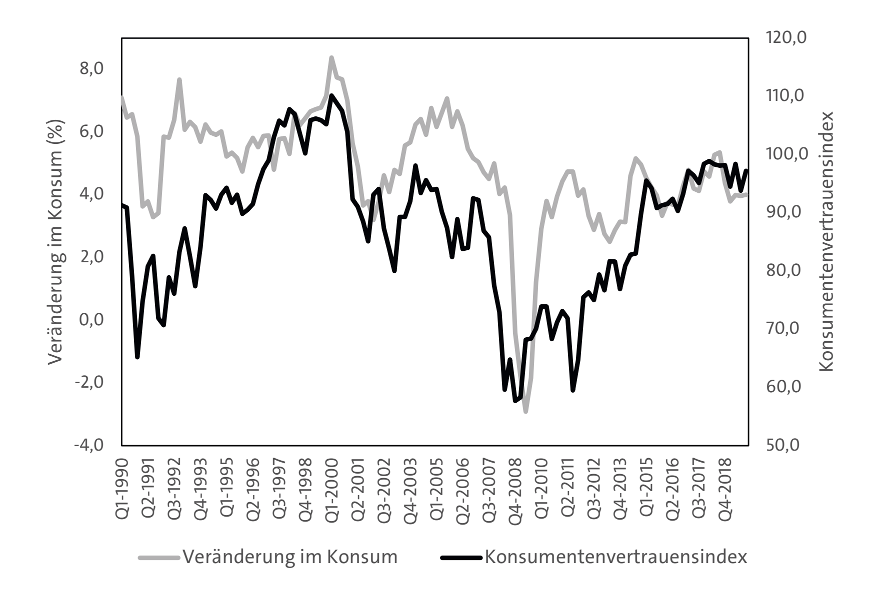


Abb. 1: Veränderungen im privaten Konsum und Konsumentenvertrauensindex in den USA
Quelle: Eigene Darstellung, basierend auf Federal State Bank of St. Louis, https://fred.stlouisfed.org/; University of Michigan http://www.sca.isr.umich.edu/tables.html.

Abbildung 1 gibt einen ersten Eindruck davon, wie sich der private Konsum und das Konsumentenvertrauen in den USA zwischen 1990 und 2019 verändert haben und inwieweit die Verläufe der beiden Variablen sich ähneln. Es ist deutlich erkennbar, dass der umfragebasierte Konsumentenvertrauensindex der University of Michigan den enormen Rückgang im privaten Konsum im Zuge der Großen Rezession früher erkennen konnte. Anders ausgedrückt lässt sich schlussfolgern, dass es möglich sein kann, mit der Analyse der Verbraucherstimmung die Veränderungen in den Konsumentscheidungen zu modellieren beziehungsweise vorauszusagen. Da allerdings die umfragebasierten Konsumentenvertrauensindizes verzögert veröffentlicht werden und ihre Ermittlung sehr zeit-und kostenintensiv ist, wurden in den letzten Jahren Daten sowohl aus den Sozialen Medien als auch Anfragen bei den Onlinesuchmaschinen verwendet, um die Stimmung der Konsumenten zeitgenauer zu messen.

Daher werden im nächsten Abschnitt die aktuellsten Studien zusammengefasst, die Daten aus den Sozialen Medien anwenden, um Konsumentscheidungen und Konsumentenvertrauen zu prognostizieren. Der nachfolgende

Abschnitt liefert eine Übersicht über die jüngsten Untersuchungen des Konsumverhaltens, die sich auf Onlinesuchanfragen stützen. Anschließend wird in dem letzten Abschnitt auf die Grenzen und Möglichkeiten der Analyse mit Onlinedaten aufmerksam gemacht und auf die möglichen Herausforderungen hingedeutet.

Soziale Medien und Konsumentscheidungen

Seit einigen Jahren werden Textdaten aus Sozialen Medien, Blogs, und Foren auch in der Erforschung der Konsumentscheidungen verwendet. Dabei werden überwiegend Kurznachrichten aus den Social-Media-Plattformen wie Twitter und Facebook benutzt. Seit 2004 stellt Facebook einen Online-Social-Networking-Dienst zur Verfügung. Mit seinen 2,7 Milliarden Nutzer:innen weltweit[2] ist Facebook die populärste Social-Networking-Plattform überhaupt. Da aber die Facebook-Posts nur durch von Nutzer:innen zugelassenen Personen gelesen oder kommentiert werden können und für Drittpersonen nicht zur Verfügung stehen, ist es nicht zielführend, ausgehend von den Posts eine allgemeine für die Konsumentscheidungen relevante Meinung ableiten zu wollen. Dies ist nur möglich wenn die Nutzer:innen in Facebook-Gruppen interagieren beziehungsweise zu Werbungen auf der Facebook-Plattform reagieren beziehungsweise Kommentare abgeben. Darüber hinaus ist es auch aus datenschutzrechtlichen Gründen nicht möglich, die Facebook-Nachrichten, die Drittpersonen nicht zur Verfügung stehen, zu analysieren. Daher konzentriert sich dieser Abschnitt auf die Kurznachrichten aus der Social-Media-Plattform Twitter.

Das Social-Networking-Unternehmen Twitter bietet seit 2006 einen Onlinedienst für Mikroblogging. Registrierte Nutzer:innen dürfen auf der Twitter-Plattform kurze Nachrichten (die sogenannten Tweets) senden und mit anderen Nutzer:innen interagieren, wobei die nicht registrierten Nutzer:innen diese Tweets nur lesen dürfen. Die Länge der Tweets ist seit November 2017 auf höchstens 280 Zeichen begrenzt, mit der Ausnahme der Tweets in chinesischer, japanischer und koreanischer Sprache, welche nur aus 140 Zeichen zusammengesetzt sein dürfen. Die 340 Millionen Twitter-Nutzer:innen weltweit senden

---

[2] Vgl. Statista, https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/, aufgerufen am 23.12.2020.

täglich durchschnittlich 500 Millionen Tweets.[3] Damit ist Twitter weltweit eine der größten Social-Media-Plattformen.

Tag für Tag steigen die Zahl der Nutzer:innen und die Zahl der Tweets an. Social-Media-Plattformen beeinflussen den Alltag in vielfacher Weise und mit steigender Tendenz. Mittlerweile ist Twitter als wichtiges Medium für Politiker:innen, Journalist:innen und Unternehmen unerlässlich geworden. Durch Twitter haben Forscher:innen Zugang zu einer riesigen Datenmenge, mit der sie „wertvolle Einblicke in das Verhalten von Menschen, ihre Stimmungen, ihre Konsummuster, ihre Sprache und ihre Abstimmungen" bekommen können. Daher wird es wichtiger, notwendiger und erkenntnisreicher denn je, Twitter-Daten zu analysieren.[4]

In letzter Zeit haben Tweets vermehrt Anwendung gefunden, um die führenden Aktienmarktindizes[5] wie Dow Jones Industrial Average, NASDAQ Composite und S&P 500 oder Rendite, Volatilität und Handelsvolumen[6] für S&P 500 vorherzusagen. Die Untersuchungsergebnisse haben gezeigt, dass sich die Vorhersagegenauigkeit durch die Einbeziehung der Stimmung in den sozialen Netzwerken verbessert.

Es gibt einige Studien, die sich mit der Prognose der Konsumentenvertrauensindikatoren anhand von Twitter-Kurznachrichten befassen[7]. Die Stimmung der Twitter-Nutzer:innen wird dabei mit den umfragebasierten Konsumentenvertrauensindizes verglichen. Um diese umfragebasierten Konsumentenvertrauensindizes zu konstruieren, wird monatlich eine zufällige Stichprobe von Verbraucher:innen befragt. Dabei beantworten die Interviewten Fragen zu

- ihrer finanziellen Situation in den letzten und in den nächsten zwölf Monaten,
- der allgemeinen wirtschaftlichen Lage in den letzten und in den nächsten zwölf Monaten,
- Preistrends in den letzten und in den nächsten zwölf Monaten,
- ihren geplanten großen Einkäufen derzeit und in den nächsten zwölf Monaten,
- Einsparungen derzeit und in den nächsten zwölf Monaten,
- Ihren Erwartungen der Arbeitslosigkeit in den nächsten 12 Monaten.

---

[3] Statista, online unter https://de.statista.com/themen/99/twitter/, zuletzt aufgerufen am 7.7.2021.

[4] Puschmann et al. (2014).

[5] Bollen/Mao/Zeng (2011); Zhang (2017); Zhang/Fuehres/Gloor (2011); Oliveira/Cortez/Areal (2017).

[6] Oliveira et al. (2017).

[7] Daas/Puts (2014); Shayaa et al. (2018); Shayaa et al. (2017).

Mit Hilfe einer Sentimentanalyse[8] wird die Stimmung der Konsumenten basierend auf ihren Kurznachrichten in den Sozialen Medien (Tweets, Facebook-Posts) gemessen.

Die Sentimentanalyse, auch Opinion Mining genannt, ist ein Forschungsgebiet, das sich mit der Analyse der Meinungen, Gefühle, Einschätzungen, Einstellungen und Emotionen der Menschen gegenüber Objekten und ihren Attributen, die sie in geschriebenem Text ausdrücken, befasst. Diese Objekte können Produkte, Dienstleistungen, Organisationen, Einzelpersonen, Veranstaltungen oder Themen sein.[9]

In Tabelle 1 werden die jüngsten Studien aus der Literatur zusammengefasst, die Textdaten überwiegend aus Sozialen Medien nutzen, um Konsumentenvertrauen beziehungsweise Konsumentscheidungen vorauszusagen.

Wie Tabelle 1 offenbart, benutzen verschiedene Studien unterschiedliche Such- und Schlagwörter, um die Kurznachrichten beziehungsweise Einträge in den Foren zu „scrapen" (das heißt, zu schürfen). Wie oben erläutert wurde, wird dabei die Sentimentanalyse eingesetzt. Durch die Korrelationsanalyse wird die Stärke des Zusammenhangs zwischen dem umfragebasierten Konsumentenvertrauensindex und dem Sentiment Score – basierend auf den Sozialen Medien – betrachtet. Ein Vergleich der Ergebnisse der verschiedenen Studien verdeutlicht, dass die jeweils gewählten Schlagwörter bei der Berechnung des Sentiments eine sehr große Rolle spielen. Die Schlagwörter, die mit den Fragen aus den Konsumentenbefragungen hergeleitet werden können, liefern eine höhere Korrelation. Außerdem ist es bei der Berechnung des Sentiments wichtig, wie der Sentiment Score definiert wird. Überwiegend verwendeten die veröffentlichten Studien bisher für die Berechnung des Sentiment-Scores, die Prozentsätze der Wörter mit negativen und positiven Sentiments, die im Textkorpus für die Sentimentanalyse zur Verfügung stehen. Als Alternative ließen sich in Zukunft Sentiment-Wörterbücher einsetzen, solange die Lexika mit den kodierten Werten für die Wörter, wie zum Beispiel Adjektive und Adverbien, zur Verfügung stehen. Diese Adjektive und Adverbien und einige für das Sentiment relevante Nomen und Verben werden dann mit Zahlen bewertet, die im Intervall von [–1, 1] liegen, um die Sentiment-Wörterbücher zu bilden.

---

[8] Näheres zur Sentimentanalyse ist in dem Beitrag von Sturm (2021) zu finden.

[9] Liu (2015).

Tab. 1: Studien mit Textdaten aus Sozialen Medien, Foren und Blogs

AUTOREN	ZIEL	METHODE	SCHLAGWÖRTER/ LAND	ERGEBNISSE
Salampasis/ Paltoglou / Giahanou (2011)	Ansichten der Konsument:innen zu Produkten und Marken analysieren	Sentimentanalyse (lexikonbasiert) – täglicher Prozentsatz der negativen, positiven oder neutralen Stimmung bei den Twitter-Daten	35 verschiedene Lebensmittelprodukte und -marken, wie „McDonalds", „milk", „seafood"	Die Stimmung der Konsument:innen gegenüber einem Produkt ändert sich kurzfristig nicht, sofern es keinen triftigen Grund gibt.
Daas / Puts (2014)	Zusammenhang zwischen dem umfrage- und Social-Media-basierten Konsumentenvertrauen in den Niederlanden	Sentimentanalyse, Korrelation, Granger-Kausalitätsanalyse, Kointegrationanalyse	Tweets, Facebook-Posts, Blogeinträge aus den Niederlanden. Schlagwörter: „economy", „job", „jobs", „consumer", „confidence", „finance", „spending", Personalpronomen auf Niederländisch und Englisch	Textbasiertes Stimmungsmaß ist hoch korreliert und kointegriert mit dem offiziellen Konsumentenvertrauensindex.
O'Connor et al. (2010)	Vergleich von Umfragewerten (für Konsumentenvertrauen und US-Wahlen) mit der Stimmung in Twitter-Daten.	Sentimentanalyse (lexikonbasiert) – positive und negative Werte der Wörter.	Konsumentenvertrauen: „economy", „job", „jobs" Wahlen: „obama, mccain" zwischen 2008 und 2009	Ab Mitte 2009 kann Stimmung aus Textnachrichten das Konsumentenvertrauen besser voraussagen als vorher.
Pekar / Binner (2017)	Modellierung von Konsumausgaben anhand von Kaufabsichten in Social Media	saisonale ARIMA, AdaBoost, Gradientenverstärkung, Word2vec	Tweets mit den Kombinationen von („I", „we"), („will", „'ll", „be going to", „be looking to", „want to", „wanna", „gonna"), („buy", „shop for", „get oneself") aus den USA	Verbesserung der Prognosegenauigkeit um 11 % bis 18 %.
Shayaa et al. (2017)	Analyse des Kaufverhaltens der Konsument:innen in Malaysia via Social Media	Sentimentanalyse, überwachtes maschinelles Lernen, Korrelationsanalyse	Kurznachrichten aus Twitter, Foren, Mainstream-Medien, Blogs, Facebook und Onlinekommentare beim Kauf von Computern und Telefonen 2015bis 2016.	schwache Korrelation zwischen der textbasierten Stimmung und dem umfragebasierten Konsumentenvertrauensindex
Shayaa et al. (2018)	Zusammenhang zwischen dem Konsumentenvertrauensindex in Malaysia und der Stimmung aus Social Media	Sentimentanalyse, Korrelationsanalyse	Tweets von 2015 bis 2019, die die Wörter „car" oder „holiday" auf Englisch und Malaysisch beinhalten.	Schwache Korrelation zwischen der Textbasierten und der umfragebasierten Stimmung.
Conrad et al. (2019)	Zusammenhang zwischen dem University of Michigan Index zum Konsumentenvertrauen (ICS) und der Stimmung aus Twitter-Daten	Korrelationsanalyse, Kointegration	Tweets vom 1. Januar 2007 bis zum 27. Juni 2014 aus den USA„ die das Wort „jobs" beinhalten	Die Ergebnisse in (O'Connor u. a. [2010]) können ab 2011 nicht repliziert werden.

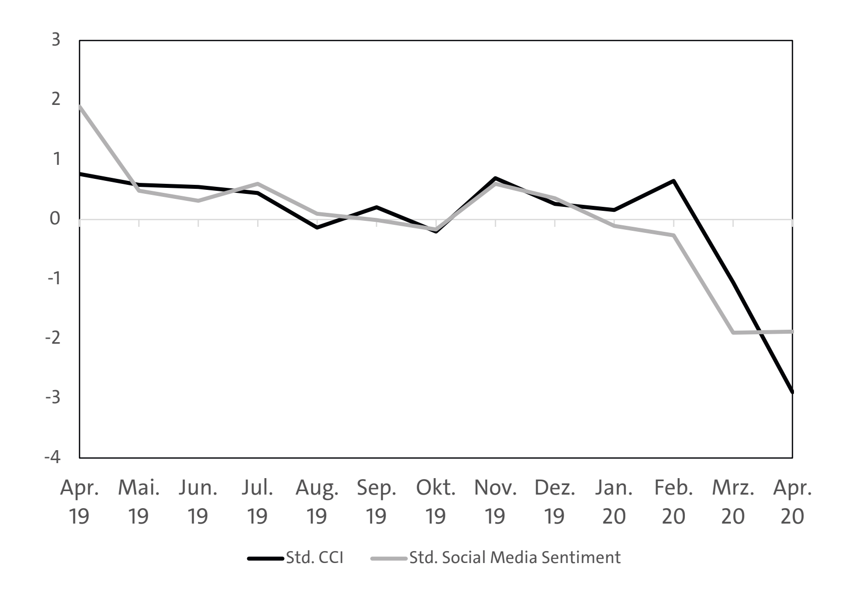


Abb. 2: Standardisierter Konsumentenvertrauensindex (CCI) in Deutschland
Quelle: https://stats.oecd.org und Standardisiertes Sentiment aus Social Media in Deutschland, April 2019 bis April 2020. Quelle: Eigene Berechnungen.

Abbildung 2 zeigt, wie gut die Bestimmung des Konsumentenvertrauens anhand von Kurznachrichten aus Social Media sein kann, wenn das Sentiment mit Hilfe eines Sentiment-Wörterbuchs berechnet wird. Ziele der Analyse, die Abbildung 2 zugrunde liegt, sind die Berechnungen des Social-Media-Sentiments der Konsument:innen und dessen Vergleich mit dem umfragebasierten Konsumentenvertrauensindex in Deutschland. Dabei wurde das Social-Media-Sentiment mit der lexikonbasierten Sentimentanalyse ermittelt. Für die Sentimentanalyse wurden die wirtschaftlich relevanten Kurznachrichten aus Twitter in der Zeit vom April 2019 bis zum April 2020 mittels dem Sentiment-Wörterbuchs SentiWS[10] für die deutsche Sprache analysiert. Insgesamt standen für diesen Zeitraum rund 502.743 Tweets zur Verfügung. Die ersten Ergebnisse haben einen Korrelationskoeffizienten von 0,83 zwischen dem Sentiment aus Social Media und dem umfragebasierten Konsumentenvertrauensindex ergeben. Dieses vielversprechende Ergebnis kann natürlich bekräftigt werden, wenn historische

[10] Remus/Quasthoff/Heyer (2010).

Tweets für Forschungszwecke zur Verfügung stehen. Standardmäßig steht leider momentan nur 1 % der gesamten Tweets kostenlos zur Verfügung, und es ist nicht möglich Tweets die älter als 9 Tage sind. Allerdings gewährt Twitter seit einigen Monaten auf Antrag für Forschungszwecke kostenlos einen vollen Zugang zum Twitter-Archiv. Mit diesem akademischen Zugang ist es möglich, pro Antrag bis zu 10 Millionen Tweets monatlich zu scrapen.

Google Trends und Konsumentenverhalten

Google Trends[11] liefert seit 2004 „einen Zeitreihenindex vom Volumen der Suchanfragen, die die Nutzer:innen in einem bestimmten geografischen Gebiet bei der Onlinesuchmaschine Google eingeben"[12]. Die Berechnung dieser Zeitreihenindizes erfolgt basierend auf den Suchanfragen je nach Bedarf jährlich, monatlich, wöchentlich und/oder täglich. Der Zeitpunkt mit den meisten Suchanfragen wird dem Wert 100 zugeordnet. Dies bedeutet wiederum, dass sich die historischen Indexwerte mit der Zeit ändern können.

Bei Google Trends werden für 605 Hauptkategorien[13] und insgesamt für 1426 Unterkategorien aggregierte Indizes, die auf den Suchanfragen basieren, veröffentlicht. Allerdings kann es vorkommen, dass nicht für jede Kategorie ein Index in einem geografischen Gebiet kalkuliert wird, da in einigen Kategorien Suchanfragen fehlen können. Darüber hinaus ist zu berücksichtigen, dass die aggregierten Indizes ohne Bereinigung der saisonalen Schwankungen berechnet werden.

Da in den letzten Jahren die Konsument:innen zunehmend ihre Einkäufe und Preisvergleiche zunehmend via Internet durchführen, können Suchanfragen für Waren und Dienste bei den Suchmaschinen eine alternative Datenquelle werden, um private Konsumausgaben zu prognostizieren.

---

[11] Vgl. https://trends.google.com/trends/.

[12] Choi/Varian (2012).

[13] Hauptkategorien sind Unter anderem Autos und Fahrzeuge, Beruf und Ausbildung, Bücher und Literatur, Computer und Elektronik, Essen und Trinken, Finanzen, Gesetz und Regierung, Gesundheit, Haus und Garten, Haustiere und wild lebende Tiere, Hobby und Freizeitbeschäftigungen, Immobilien, Internet und Kommunikation, Kunst und Unterhaltung, Mensch und Gesellschaft, Nachrichten, Naturwissenschaften, Online-Communitys, Referenz, Reisen, Schönheit und Fitness, Shopping, Spiele, Sport, Unternehmen und Industrie.

Neben der Erstellung von Prognosen können Daten aus Google Trends auch helfen, die Gegenwart besser zu verstehen.[14] Zum Beispiel ist es mit Hilfe der Daten aus Suchmaschinen möglich, unter anderem Influenzawellen[15], Inflationserwartungen[16], Arbeitslosigkeitsrate[17] beziehungsweise Anträge auf Arbeitslosigkeit[18], Autoverkäufe[19], Hausverkäufe[20], Zahl der Museumsbesucher[21], Hauspreise[22], Einnahmen von Kinofilmen[23] und Bruttoinlandsprodukt[24] zu prognostizieren.

Im vorliegenden Beitrag liegt das Hauptaugenmerk auf dem Einsatz von Google-Trends-Daten um Konsumentenvertrauen- und Entscheidungen sowie privaten Konsum vorauszusagen. Um den privaten Konsum in den USA zu prognostizieren,[25] wurden 56 wirtschaftlich relevante Google-Trends-Kategorien herauskristallisiert. Die Vorhersagemodelle basieren dabei auf der Erkenntnis, dass der Konsumentenvertrauensindex den privaten Konsum in den USA erklären kann.[26] Anhand einer Faktorenanalyse wird die Zahl der Kategorien reduziert und die geschätzten Faktoren fließen anschließend in die Vorhersagemodelle ein. Die Ergebnisse zeigen, dass die Modelle, in denen die Faktoren aus den Google-Trends-Daten als erklärende Variable zusammen mit den für die Konjunktur relevanten makroökonomischen Variablen (das heißt reales persönliches Einkommen, S&P-500-Aktienmarktindex, Dreimonatszinssatz) auftauchen, bessere Prognosen liefern als die Baseline-Modelle. Hier muss auch darauf hingewiesen werden, dass die Studie den Zeitraum zwischen 2005 und 2009 betrachtet hat.

---

[14] Choi/Varian (2009b); Choi/Varian (2009a), (2009b), (2012).

[15] Ginsberg et al. (2009).

[16] Guzmán (2011).

[17] D'Amuri/Marcucci (2009); Askitas/Zimmermann (2009).

[18] Choi/Varian (2009a).

[19] Choi/Varian (2009b); Choi/Varian (2012); Carrière-Swallow/Labbé (2013).

[20] Choi/Varian (2009b); Choi/Varian (2012).

[21] Volchek et al. (2018).

[22] Wu/Brynjolfsson (2015).

[23] Asur/Huberman (2010).

[24] Liu/Xu/Fan (2018); Qian (2018).

[25] Vosen/Schmidt (2011).

[26] Wilcox (2007), Ludvigson (2004); Carroll/Fuhrer/Wilcox (1994).

In einer anderen Studie zur Prognose des privaten Konsums in den USA wurden neben den finanziellen Variablen wie S&P-500-Aktienmarktindex oder kurzfristige (Dreimonats) und langfristige (Zehnjahres)Zinssätze lediglich 25 Google-Trends-Kategorien benutzt.[27] Obwohl diese Studie nur den Zeitraum zwischen 2007 und 2009 analysiert, haben die Ergebnisse mit den Faktoren aus den Google-Trends-Kategorien wieder eine bessere Voraussage geliefert.

Ähnliche Ergebnisse mit Google-Trends-Daten wurden in einer anderen Studie auch für den Konsum in Deutschland erzielt.[28] Dabei wurden lediglich 41 Google-Trends-Kategorien benutzt und die Daten aus der Zeit von 2004 bis 2010 analysiert. Basierend auf diesen Erkenntnissen wird beim RWI – Leibniz-Institut für Wirtschaftsforschung der RWI-Konsumindikator[29] berechnet. Dem Institut zufolge weist der „RWI-Konsumindikator gegenüber dem privaten Verbrauch einen Vorlauf von einem Quartal“[30] auf.

In einer aktuelleren Studie[31] fungierten 66 konsumrelevante Google-Trends-Kategorien zusammen mit den nachrichtenrelevanten Google-Trends-Daten[32] als erklärende Variablen, um die Veränderung des privaten Konsums in den USA zu prognostizieren beziehungsweise in Echtzeit vorauszusagen (Nowcast[33]). Dabei wurden „recession“ und „layoff“ als Suchwörter für die nachrichtenrelevanten Google-Trends-Daten benutzt. Die Analyse, basierend auf den monatlichen Daten für den Zeitraum zwischen 2009 und 2014, hat gezeigt, dass die nachrichtenrelevanten Google-Trends-Daten für die Veränderung des Konsums von Gebrauchsgütern nützlich sein können und die Modelle mit den konsumrelevanten Google-Trends-Kategorien Konsumabsichten erklären können. Darüber hinaus können konsum- und nachrichtenrelevante[34] Google-Trends-Daten zusammen die Prognosegüte verbessern und dabei helfen, den Konsum einen Monat im Voraus zu prognostizieren.

In einer Studie mit zwei Google-Trends-Kategorien (das heißt „motorcycle“, „birth control“) ließ sich die Voraussage des Konsumentenvertrauensindex

---

[27] Kholodilin/Podstawski/Siliverstovs (2010).

[28] Vosen/Schmidt (2012).

[29] RWI (2010), S.103. Online abrufbar unter https://www.rwi-essen.de/forschung-und-beratung/wachstum-konjunktur-oeffentliche-finanzen/projekte/221/.

[30] RWI-Konsumindikator, https://www.rwi-essen.de/konsumindikator

[31] Woo/Owen (2019).

[32] Google Trends, https://trends.google.de/trends/explore?cat=16&date=all_2008&gprop=news.

[33] Näheres zu „Nowcasts“ ist in dem Beitrag von Maaß (2021) zu finden.

[34] Google Nwes, https://news.google.com.

in Sri Lanka im Vergleich mit einem autoregressiven Prozess erster Ordnung um 5,2 % verbessern.[35] Diese zwei Google-Trends-Kategorien wurden mit Hilfe der Spike-und-Slab-Regression unter Berücksichtigung der insgesamt 1104 Google-Trends-Kategorien für die Zeit vom September 2012 bis zum April 2017 bestimmt.

Google-Trends-Daten werden auch verwendet, um Konsumentenentscheidungen und -verhalten für Mode vorauszusagen. In einer Studie wird zum Beispiel Google-Trends-Serie für die Suchanfrage „Burberry" mit verschiedenen Modellen geschätzt.[36] Die Prognosegüte des bereinigten („enoised") neuronalen Netzwerk-Modells (DNNAR) wird dabei mit den Prognosen, basierend auf den konventionellen Zeitreihenmodellen wie ARIMA, exponentielle Glättung oder ARMA mit saisonalen Trends, verglichen. Der Vergleich ergibt, dass das DNNAR-Modell bessere Ergebnisse liefert als die anderen Modelle. Allerdings weisen die Autoren darauf hin, dass die Google-Trends-Daten für die moderelevanten Suchanfragen hohe saisonale Schwankungen zeigen, wodurch die Prognosen für die moderelevanten Suchanfragen mit wenigen saisonalen Schwankungen schwer vorherzusagen wären.

Schlussfolgerungen

Textdaten aus Sozialen Medien bergen große Chancen und Herausforderungen für die Praxis. Die ersten vielversprechenden Ergebnisse zeigen, dass die Sentiment Scores, die auf der Basis von Kurznachrichten aus Sozialen Medien berechnet worden sind, in der Zukunft die umfragebasierten Konsumentenvertrauensindizes ersetzen können. Da die Durchführung und Analyse von Umfragen viel Zeit und Geld kostet, können durch die Nutzung von Textdaten aus Sozialen Medien viel Zeit und Ressourcen gespart werden. Eine der größten Herausforderungen dabei ist die Auswahl der richtigen Such- und Schlagwörter für die Praxis. Eine mögliche Lösung dafür könnte der Einsatz von Algorithmen beziehungsweise künstlicher Intelligenz sein. Dabei müssen die Algorithmen mit Daten zunächst trainiert werden. Mit der Zeit können die Algorithmen durch den Lerneffekt verbesserte Prognosen liefern.

Die Nützlichkeit von Google-Trends-Daten unter anderem zu Voraussage von Konsumentscheidungen hat sich schon längst erwiesen. Dennoch könnten

---

[35] Zhang (2017).

[36] Silva et al. (2019).

die Prognosen verbessert werden, indem die Daten von Onlinesuchanfragen und Textdaten aus Sozialen Medien zusammen in die Analyse einfließen. Es gibt schon einige erste Versuche in dieser Richtung.[37] Um beispielsweise die Autoverkäufe in Deutschland vorauszusagen, wird das Suchverhalten der Verbraucher:innen mit der Stimmung der Verbraucher:innen in Bezug auf eine Marke oder ein Produkt kombiniert.[38] Die Ergebnisse der Analyse haben gezeigt, dass die Modelle mit Google-Trends- und Twitter-Daten bessere Prognosen liefern als die Modelle, die nur Suchanfragedaten oder Textdaten aus Sozialen Medien verwenden.

Es muss hier auch noch einmal betont werden, dass Google-Trends-Daten sich mit der Zeit ändern können – mit entsprechenden Rückwirkungen auf die Berechnung von Indexwerten für einen bestimmten Zeitpunkt. Dies wiederum macht die Replizierbarkeit der Studien schwierig. Darüber hinaus wurden in den Studien wegen der hohen Saisonalität, die Google-Trends-Daten in Wachstumsraten umgewandelt, um saisonale Schwankungen zu eliminieren. Da mittlerweile längere Zeitreihen zur Verfügung stehen, wird es bald möglich sein, Google-Trends-Daten mit den konventionellen Methoden wie X12-ARIMA oder X-13-ARIMA-SEATS saisonal zu bereinigen. Dies wird eine einfachere Interpretation der Ergebnisse ermöglichen.

Die Wahl der richtigen Schlagwörter und Kategorien ist für die Analyse mit Google-Trends-Daten sehr wichtig. In einigen Studien wurde vorgeschlagen, Google Correlate zu nutzen, um die richtigen Schlagwörter zu finden.[39] Bei Google Correlate ist es möglich, anhand einer Zeitreihe einige relevante Google-Trends-Kategorien oder Unterkategorien zu bekommen. Diese vorgeschlagenen Google-Trends-Kategorien haben schlussendlich einen ähnlichen Verlauf wie die Zeitreihe, die analysiert werden soll. Allerdings können sich dabei Suchwörter ergeben, zu denen lediglich ein begrenzter ein Zusammenhang hergestellt werden kann. Daher ist es wichtig, die Ergebnisse herauszufiltern und mit Vorsicht zu interpretieren.

Eine weitere Möglichkeit wäre die Nutzung von Google-Analytics-Daten, die keine Indexdaten sind. Mit deren Hilfe wäre es möglich, replizierbare Analysen durchzuführen.

---

[37] Skodda/Benthaus (2015).

[38] Ebenda.

[39] Sawaengsuksant (2018).

Literatur

Askitas, Nikolaos und Zimmermann, Klaus F. (2009): Google Econometrics and Unemployment Forecasting. In: Applied Economics Quarterly, Vol. 55 (2), S. 107–120, DOI https://doi.org/10.3790/aeq.55.2.107.

Asur, Sitaram und Huberman, Bernardo A. (2010): Predicting the Future with Social Media. In: Hoeber, Orland (Hrsg.): IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, S. 492–499, DOI https://doi.org/10.1109/WI-IAT.2010.63.

Bollen, Johan; Mao, Huina; Zeng, Xiaojun (2011): Twitter Mood Predicts the Stock Market. In: Journal of Computational Science, Vol. 2 (1), S. 1–8, DOI https://doi.org/10.1016/j.jocs.2010.12.007.

Carrière-Swallow, Yan und Labbé, Felipe (2013): Nowcasting with Google Trends in an Emerging Market. In: Journal of Forecasting, Vol. 32 (4), S. 289–298, DOI https://doi.org/10.1002/for.1252.

Carroll, Christopher D.; Fuhrer, Jeffrey C.; Wilcox, David (1994): Does Consumer Sentiment Forecast Household Spending? If so, why? In: The American Economic Review, Vol. 84 (5), S. 1397–1408. Online abrufbar unter https://www.jstor.org/stable/2117779.

Choi, Hyunyoung und Varian, Hal (2009a): Predicting Initial Claims for Unemployment Insurance Using Google Trends. Technical report, Google, Inc. Online abrufbar unter http://research.google.com/archive/papers/initialclaimsUS.pdf.

Choi, Hyunyoung und Varian, Hal (2009b): Predicting the Present with Google Trends. Technical Report, Google, Inc., DOI https://doi.org/10.2139/ssrn.1659302.

Choi, Hyunyoung und Varian, Hal (2012): Predicting the Present with Google Trends. In: Economic Record, Vol. 88 (1), S. 2–9, DOI https://doi.org/10.1111/j.1475-4932.2012.00809.x.

Conrad, Frederick G.; Gagnon-Bartsch, Johann A.; Ferg, Robyn A.; Schober, Michael F.; Pasek, Josh; Hou, Elizabeth (2019): Social Media as an Alternative to Surveys of Opinions About the Economy. In: Social Science Computer Review, DOI https://doi.org/10.1177/0894439319875692.

D'Amuri, Francesco und Marcucci, Juri (2009): „Google it!“ Forecasting the US unemployment rate with a Google job search index. MPRA Paper No. 18248, DOI http://dx.doi.org/10.2139/ssrn.1594132.

Daas, Piet J. H. und Puts, Marco (2014): Social Media Sentiment and Consumer Confidence. Statistics Papers Series, No. 5, Europäische Zentralbank, Frankfurt am Main, DOI https://doi.org/10.2866/11606.

Ginsberg, Jeremy; Mohebbi, Matthew H.; Patel, Rajan S.; Brammer, Lynnette; Smolinski, Mark S.; Brilliant, Larry (2009): Influenza Epidemics Using Search Engine Query Data. In: Nature, Vol. 457 (7232), S. 1012–1014, DOI https://doi.org/10.1038/nature07634.

Guzmán, Giselle (2011): Internet Search Behaviour as an Economic Forecasting Tool: The Case of Inflation Expectations. In: Journal of Economic and Social Measurement, Vol. 36 (3), S. 119–167, DOI https://doi.org/10.3233/JEM-2011-0342.

Kholodilin, Konstantin A.; Podstawski, Maximilian; Siliverstovs, Boriss (2010): Do Google Searches Help in Nowcasting Private Consumption? A Real-Time Evidence for the US. In: KOF Swiss Economic Institute Working Paper No. 256, ETH Zürich, DOI http://dx.doi.org/10.2139/ssrn.1615453.

Liu, Bing (2016): Sentiment Analysis. Mining, Opinions, Sentiments, and Emotions. Cambridge University Press, Cambridge, UK, DOI https://doi.org/10.1017/CBO9781139084789.

Liu, Taoxiong; Xu, Xiaofei; Fan, Fangda (2018): Forecasting Chinese GDP Using Online Data. In: Emerging Markets Finance and Trade, Vol. 54 (4), S. 733–746, DOI https://doi.org/10.1080/1540496X.2016.1216841.

Ludvigson, Sydney C. (2004): Consumer Confidence and Consumer Spending. In: Journal of Economic Perspectives, Vol. 18 (2), S. 29–50, DOI https://doi.org/10.1257/0895330041371222.

Maaß, Christina (2021): Nowcast als Forecast. Neue Verfahren der BIP Prognose in Echtzeit. In: Straubhaar, Thomas (Hrsg.): Neuvermessung der Datenökonomie, Hamburg University Press, Hamburg, S. 96.

O'Connor, Brendan T.; Balasubramanyan, Ramnath; Routledge, Bryan R.; Smith, Noah A. (2010): From Tweets to Pools: Linking Test Sentiment to Public Opinion Time Series. In: AAAI Publications, Fourth International Conference on Weblogs and Social Media, Palo Alto.

Oliveira, Nuno; Cortez, Paulo; Areal, Nelson (2017): The Impact of Microblogging Data for Stock Market Prediction: Using Twitter to Predict Returns, Volatility, Trading Volume and Survey Sentiment Indices. In: Expert Systems with Applications, Vol. 73, S. 125–144, DOI https://doi.org/10.1016/j.eswa.2016.12.036.

Pekar, Viktor und Binner, Jane (2017): Forecasting Consumer Spending from Purchase Intentions Expressed on Social Media. In: Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Association for Computational Linguistics, Stroudsburg, PA, DOI https://doi.org/10.18653/v1/W17-5212.

Puschmann, Cornelius; Bruns, Axel; Mahrt, Merja; Weller, Katrin; Burgess, Jean (2014): Epilogue Why Study Twitter? In: Weller, Katrin; Bruns, Axel; Burgess, Jean; Mahrt, Merja; Puschmann, Cornelius (Hrsg.): Twitter and Society. Peter Lang, New York, S. 425–432, DOI https://doi.org/10.3726/978-1-4539-1170-9.

Qian, Eric (2018): Nowcasting Indian GDP with Google Search Data. Online abrufbar unter https://core.ac.uk/reader/210594347.

Remus, Robert; Quasthoff, Uwe; Heyer, Gerhard (2010): SentiWS – A Publicly Available German-language Resource for Sentiment Analysis. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta.

RWI (2010): Analyse und Prognose des Spar- und Konsumverhaltens privater Haushalte. RWI Projektberichte, URN http://nbn-resolving.de/urn:nbn:de:hbz:061:3-15339, <zuletzt aufgerufen am 3.4.2021>.

Salampasis, Michail; Paltoglou, Giorgos; Giahanou, Anastasia (2011): Using Social Media for Continuous Monitoring and Mining of Consumer Behaviour. In: Salampasis, Michail; Matopoulos, Aristides (Hrsg.): Proceedings of the International Conference on Information and Communication Technologies for Sustainable Agri-production and Environment (HAICTA 2011), S. 192–202, DOI https://doi.org/10.1504/IJEB.2014.057905.

Sawaengsuksant, Paphatsorn (2018): Standardized Approach in Developing Economic Indicators using Internet Searching Applications. IFC – Bank Indonesia International Workshop and Seminar on „Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data".

Shayaa, Shahid; Ainin, Sulaiman; Jaafar, Noor Ismawati; Zakaria, Shamsul Bahri; Phoong, Seuk Wai; Yeong, Wai Chung; Al-Garadi, Mohammed Ali; Muhammad, Ashraf; Zahid Piprani, Arsalan (2018): Linking Consumer Confidence Index and Social Media Sentiment Analysis. In: Cogent Business & Management, Vol. 5 (1), S. 1–12, DOI https://doi.org/10.1080/23311975.2018.1509424.

Shayaa, Shahid; Al-Garadi, Mohammed Ali; Piprani, Arsalan Zahid; Ashraf, Muhammad; Sulaiman, Ainin (2017): Social Media Sentiment Analysis of Consumer Purchasing Behavior vs Consumer Confidence Index. In: Proceedings of the International Conference on Big Data and Internet of Thing – BDIOT2017. New York, 2017, S. 32–35, DOI https://doi.org/10.1145/3175684.3175712.

Silva, Emmanuel; Hassani, Hossein; Madsen, Dag; Gee, Liz (2019): Googling Fashion: Forecasting Fashion Consumer Behaviour Using Google Trends. In: Social Sciences, Vol. 8 (4), S. 111, DOI https://doi.org/10.3390/socsci8040111.

Skodda, Christoph und Benthaus, Janek (2015): Investigating Consumer Information Search Behavior and Consumer Emotions to Improve Sales Forecasting. In: Proceedings of the 21st American Conference on Information Systems (AMCIS).

Sturm, Silke (2021): Textdaten. Anwendungen und Herausforderungen. In: Straubhaar, Thomas (Hrsg.): Neuvermessung der Datenökonomie, Hamburg University Press, Hamburg, S. 126.

Volchek, E.; Song, H.; Law, R.; Buhalis, D. (2018): Forecasting London Museum Visitors Using Google Trends Data. In: e-Review of Tourism Research. Online abrufbar unter https://journals.tdl.org/ertr/index.php/ertr/article/view/121.

Vosen, Simeon und Schmidt, Torsten (2012): A Monthly Consumption Indicator for Germany based on Internet Search Query Data. In: Applied Economics Letters, Vol. 19 (7), S. 683–687, DOI https://doi.org/10.1080/13504851.2011.595673.

Vosen, Simeon und Schmidt, Torsten (2011): Forecasting Private Consumption: Survey-based Indicators vs. Google Trends. In: Journal of Forecasting, Vol. 30 (6), S. 565–578, DOI https://doi.org/10.1002/for.1213.

Vuchelen, Jef (2004): Consumer Sentiment and Macroeconomic Forecasts. In: Journal of Economic Psychology, Vol. 25 (4), S. 493–506, DOI https://doi.org/10.1016/S0167-4870(03)00031-X.

Wilcox, James A. (2007): Forecasting Components of Consumption with Components of Consumer Sentiment. In: Business Economics, Vol. 42 (4), S. 22–32, DOI https://doi.org/10.2145/20070403.

Woo, Jaemin und Owen, Ann L. (2019): Forecasting Private Consumption with Google Trends Data. In: Journal of Forecasting, Vol. 38 (2), S. 81–91, DOI https://doi.org/10.1002/for.2559.

Wu, Lynn und Brynjolfsson, Erik (2015): The Future of Prediction. In: Goldfarb, Avi; Greenstein, Shane M.; Tucker, Catherine E. (Hrsg.): Economic Analysis of the Digital Economy. National Bureau of Economic Research, Cambridge, MA, S. 89–118, DOI https://doi.org/10.7208/9780226206981.

Zhang, Haiyang (2017): Incorporating Google Trends Data in Predicting Consumer Confidence in Sri Lanka. Online abrufbar unter https://ssrn.com/abstract=3085010.

Zhang, Xue; Fuehres, Hauke; Gloor, Peter A. (2011): Predicting Stock Market Indicators Through Twitter „I hope it is not as bad as I fear". In: Procedia – Social and Behavioral Sciences, Vol. 26, S. 55–62, DOI https://doi.org/10.1016/j.sbspro.2011.10.562.

Implizite Motive in der politischen Kommunikation

Niklas Scheffer, Silke Sturm und Zahurul Islam

Einleitung

Lange galt in der politischen Ökonomie der Homo Oeconomicus als theoretischer Ausgangspunkt. Demnach orientieren sich Wähler:innen bei ihrer Stimmabgabe an der ideologischen Distanz zwischen sich selbst und den Parteien. Der geringste Abstand im euklidischen Raum entscheidet über die Allokation der Stimmen und den Wahlerfolg. Psychologische Beiträge zeigen, dass die Rationalitätsannahmen des Homo Oeconomicus gerade in der politischen Entscheidungsbildung nicht haltbar sind. So stellen etwa die Darstellung und Präsentation von Informationen einen relevanten Faktor in der Entscheidungsfindung dar. Schnellenbach und Schuber stellen die wachsende Bedeutung verhaltensökonomischer und psychologischer Ansätze in der politischen Ökonomie heraus.[1] Rein rationale Kosten-Nutzen-Erwägungen können das Kommunikations- und Wahlverhalten nicht erklären, emotionale oder scheinbar „irrationale“ Bestandteile scheinen einen erheblichen Einfluss auf Wähler:innen und Politiker:innen zu haben.

Der vorliegende Beitrag konzentriert sich auf diesen unbewussteren Aspekt. Soziale Medien bieten die einmalige Möglichkeit, die Interaktion von Parteien und Wähler:innen zu analysieren. Es werden mit über maschinelles Lernen trainierten, Algorithmen mehr als 30.0000 Facebook-Posts der im deutschen Bundestag vertretenen Parteien hinsichtlich der drei grundlegenden emotionalen beziehungsweise impliziten Motivdimensionen Macht, Bindung und Leistung analysiert. Der Einbezug von Motivmustern erlaubt es, konkrete Rückschlüsse auf eine politische Orientierung der Handelnden zu treffen, die rational nicht erklär-

---

[1] Schnellenbach/Schubert (2015).

bar erscheint, sondern eher emotional bedingt wirkt. So haben Studien von Winter[2] ergeben, dass ein bestimmtes Motivmuster – Macht hoch und Bindung tief – Krisen oder sogar kriegerische Auseinandersetzungen, die allen rationalen Kosten-Nutzen-Erwägungen widersprechen, ankündigt und damit prognostizierbar macht. Es ist von Erkenntnisinteresse, wie Parteien und Politiker:innen mit ihren Wähler:innen abseits rationaler Logik kommunizieren und ob sich emotional verankerte Motivmuster bestimmten Parteien zuordnen lassen.

Neben den Motivdimensionen werden die Mitteilungen nach den grundlegenden Themenschwerpunkten innerparteilicher Kommunikation und politikrelevanter Kommunikation auf der Basis eines unüberwachten Topic-Modells unterschieden. Ziel der Analyse ist es, durch die Kombination der Messmethoden zu interpretierbaren und für weitere Untersuchungen relevanten Ergebnissen zu kommen.

Im folgenden Beitrag werden die „großen drei" impliziten Motive dargestellt und ihre Messung erläutert. Danach werden der Algorithmus und das maschinelle Lernen vorgestellt. Im nachfolgenden Abschnitt werden die zugrundeliegenden Daten und die Hypothesen beschrieben. Der anschließende Abschnitt widmet sich den Interpretationsmöglichkeiten und der politischen Relevanz der Ergebnisse. Der Beitrag schließt mit einer Diskussion.

Implizite Motive

Die „großen drei" impliziten Motive

Bereits seit den 1950er-Jahren ist Psycholog:innen bekannt, dass das menschliche Gehirn über mehrere Systeme zum Wahrnehmen, Verarbeiten und Speichern von Informationen verfügt, von denen nur ein Bruchteil bewusst von uns wahrgenommen wird.[3] Implizite Motive spielen eine große Rolle dabei, wie unser Verhalten initiiert, aufrechterhalten und gestoppt wird.[4] Sie orientieren, wählen und aktivieren unser Verhalten vor allem durch emotionale, in alten Teilen des Gehirns verankerte Impulse.[5]

---

[2] Winter (1993); Winter (2004); Winter (2007).

[3] Schultheiss/Brunstein (2010), S. xv.

[4] McClelland (1999).

[5] Schultheiss/Brunstein (2010).

Ein Motiv lässt sich als eine stabile individuelle Disposition definieren, auf bestimmte Anreize in unserer Umwelt emotional besonders stark zu reagieren.[6] Eine Person mit einem besonders ausgeprägten Motiv richtet ihr Verhalten danach aus, bestimmte Reize direkt zu suchen (oder zu vermeiden), um durch ihre Reaktion auf sie Befriedigung zu erreichen.[7] Kein Motiv operiert dabei über die gesamte Zeit, sondern wird immer nur durch bestimmte Anreize oder Ziele aktiviert. Individuen unterscheiden sich aber darin, wie schnell sie durch ein bestimmtes Motiv angesprochen werden. Motive sind somit sowohl stabile individuelle Persönlichkeitsmerkmale (traits) als auch zeitlich und situationsübergreifend variabel (states).[8]

Die genaue Zahl der Motive ist umstritten. Während Murrays Motivkatalog von 1938 noch fast 20 Motive umfasste, geht heute ein großer Teil der Motivforschung von drei grundlegenden Dimensionen menschlicher Motivation aus: Macht, Bindung und Leistung. Das Machtmotiv wird als das Bedürfnis definiert, andere zu beeinflussen, zu kontrollieren oder Wirkung auf eine Person, Gruppe oder die ganze Welt zu erzielen.[9] Das Bindungsmotiv wird als Bedürfnis definiert, eine positive Beziehung zu einer anderen Person oder Gruppe aufzubauen aufrechtzuerhalten oder wiederherzustellen.[10] Das Leistungsmotiv definiert sich, durch das Bedürfnis, etwas besser machen zu wollen oder einen gewissen Leistungsstandard zu erfüllen.[11]

Neben diesen drei Motivdimensionen gibt es auch Bedürfnisse wie Hunger, Durst, Hygiene, Sex, Durchsetzung, Integration, Sicherheit und Erkenntnis, die teilweise aber starke Überlappungen mit den „großen drei“ Motiven aufweisen. Die Reduktion auf drei grundlegende Motive dient daher vor allem dem Zweck, mit möglichst wenigen Motiven Verhalten und dessen Änderungen möglichst vollständig erklären zu können.[12] Wie der größte Teil der Motivforschung legt sich auch der Datensatz, der diesem Beitrag zugrunde liegt, auf diese drei Basismotive fest.

---

[6] Winter (2007).

[7] Schultheiss/Brunstein (2010).

[8] Winter (2007).

[9] Ebenda.

[10] Pang (2010).

[11] McClelland (1999).

[12] McClelland (1999); Schultheiss/Brunstein (2010).

Messung der impliziten Motive

Der erste Psychologe, der den Versuch unternahm, die Motivforschung zu systematisieren, war 1935 Henry A. Murray. Angelehnt an die Freud'sche Methode der freien Assoziation, entwickelte er den Thematic Apperception Test (TAT).[13] Murray ersetzte die spontaneren und unstrukturierteren Ergebnisse, die Freud durch seine Patient:innen erhielt, durch einen standardisierten Satz mehrdeutiger Bilder. Jedes der 20 Bilder sollte ein bestimmtes Motiv hervorrufen. Nach dem Betrachten der Bilder wurden die Probanden aufgefordert, sich kurze Geschichten auszudenken und diese niederzuschreiben. Die Geschichten wurden anschließend hinsichtlich einer feststehenden Anzahl von beinahe 20 Motiven ausgewertet.[14]

Auf Atkinson und McClelland geht die erste streng empirisch konzipierte Studie in der Motivforschung zurück.[15] Bei einer Versuchsgruppe wird zunächst gezielt ein bestimmtes Motiv angeregt, danach wird die Gruppe aufgefordert, das TAT-Testformular auszufüllen. Die TAT-Ergebnisse der aktivierten Versuchsgruppe werden anschließend mit den Ergebnissen einer Kontrollgruppe verglichen, bei der das Motiv nicht angeregt wurde. Mit dieser Methode legten Atkinson und McClelland die Grundlage für die Messung von Motiven. Die Ergebnisse konnten mit verschiedenen Studien nach dem gleichen Studiendesign bestätigt werden.[16] In späteren Studien zeigte sich, dass die Motive zwar situativ variieren, aber grundsätzlich zeitstabil in den TAT-Ergebnissen zu erkennen sind. Schlussfolgerung ist, dass es sich bei Motiven um zeitübergreifende Persönlichkeitsmerkmale handelt.[17]

Die Analysen wurden in der zweiten Hälfte des 20. Jahrhunderts auf eine kollektive, gesellschaftliche Ebene angewendet.[18] Es konnte dabei gezeigt werden, dass die „großen drei“ Motive auch auf kollektiver Ebene Erklärungswert besitzen. Für das Leistungsmotiv konnte etwa ein Zusammenhang zwischen leistungsbezogenen Inhalten in Schulbüchern und späterem Wirtschaftswachstum nachgewiesen werden.[19] Winter konnte für die Zeit zwischen 1603 und 1988

---

[13] Murray (1943); Weinberger/Cotler, Tanya, Fishman, Daniel (2010).

[14] McClelland (1999).

[15] McClelland/Atkinson (1948).

[16] Shipley Jr. / Veroff (1952); Steele (1977).

[17] Fodor (2010).

[18] McClelland (1961).

[19] Ebenda.

ein kollektives Motivmuster, hohes Machtmotiv und niedriges Bindungsmotiv, in Großbritannien jeweils ein Jahr vor Kriegseintritt nachweisen.[20] Für politische Krisen lässt sich das gleiche Motiv-Inhaltsmuster für die Kommunikation politischer Entscheidungsträger:innen zeigen.[21]

Die Ergebnisse zeigen eine breite Anwendbarkeit der Motivforschung. Weitere Studien legen überzeugend die Möglichkeit dar, über inhaltsanalytische Methoden konkrete Verhaltensmuster von Individuen und Kollektiven zu prognostizieren.

Politische Ökonomie

Der Ausgangspunkt des rationalen Wählers nach Downs[22], der unter der Rationalitätsannahme und der Annahme vollständiger Information, eine seinen Nutzen maximierende Entscheidung trifft, führt im eindimensionalen Raum zur Konvergenz politischer Parteien.[23] Die diesem Modell zugrundeliegenden Annahmen von Invarianz, Vollständigkeit aller Informationen, Eindimensionalität des Politikraums und dem Eigennutzaxiom wurden im Laufe der Zeit stark kritisiert.[24]

Kahneman und Tversky/Kahneman zeigen, dass Framing von identischen Informationen, also die Kontextualisierung, die Entscheidungen von Agent:innen beeinflusst.[25] Bezogen auf den rationalen Wähler können die strengen Rationalitätsannahmen widerlegt werden. Menschen unterliegen kognitiven Verzerrungen, die ihre Entscheidungsfindung unbewusst beeinflussen.[26]

Cowen kritisiert die Axiome des Downs‘schen Wählers von Rationalität und vollkommenen Informationen aus einer verhaltensökonomischen Perspektive.[27] Er argumentiert, dass jeder Mensch vorgefestigte Überzeugungen („biases“) in sich trägt, welche ihn unbewusst daran hindern, rationale Entscheidungen zu treffen. Die meisten Menschen sind davon überzeugt, die Wahrheit

---

[20] Winter (1993).

[21] Winter (2004); Winter (2007).

[22] Downs (1957).

[23] Herz/Weinberger (2019).

[24] Bénabou/Tirole (2016); Caplan (2011); Cowen (2005); Kahneman (2003); Tversky/Kahneman (1986).

[25] Kahneman (2003); Tversky/Kahneman (1986).

[26] Simon/Fassnacht (1982).

[27] Cowen (2005).

zu kennen, und hinterfragen diese auch dann nicht, wenn ihnen Personen mit einer offensichtlich größeren Expertise, dieser widersprechen.[28] Erhalten Menschen Informationen, die nicht mit ihrem Selbstbild übereinstimmen, so verwerfen sie diese oder interpretieren sie so, dass ihre vorherigen Glaubenssätze sogar noch gestärkt werden.[29] Cowen schließt aus diesen Erkenntnissen, dass politische Kommunikation und Informationsaustausch in Demokratien, entgegen der Annahme von Downs (1957), nicht zu einer Konvergenz der Meinungen, sondern zu einer vertieften Polarisierung tendiert.[30]

Bénabou und Tirole weisen darauf hin, dass kognitive Verzerrungen und systematische Selbsttäuschung wichtige psychologische und funktionale Bedürfnisse des Einzelnen erfüllen.[31] Beispiele sind das Vertrauen in die eigenen Fähigkeiten, das moralische Selbstwertgefühl, die Verringerung von Angst, die soziale Identität, die politische Ideologie und der religiöse Glaube. Dieser Befund wird auch von der Forschung unterstützt. So haben mehrere Studien ergeben, dass sich selbstüberschätzende und selbsttäuschende Personen glücklicher sind als Personen, die ein weniger hohes Maß an Selbsttäuschung aufweisen. In der Tat ist eine hohe Objektivität bei der Verarbeitung von Informationen sogar eher ein Anzeichen für eine Depression.[32]

Zusammengefasst sind verhaltensökonomische Ansätze in der Public-Choice-Literatur zunehmend relevant. Die im vorherigen Abschnitt angesprochenen impliziten Motive wurden im politischen Kontext von Winter nachgewiesen.[33] Sie lassen sich als intra- und interindividuell wirksame „frames" beziehungsweise „biases" interpretieren. Die Motive werden im vorliegenden Beitrag zur Erklärung angebotsseitiger politischer Kommunikation verwendet. Die Herangehensweise trägt eine neue verhaltensökonomische Komponente zur Erklärung des Verhaltens politischer Akteur:innen bei.

---

[28] Caplan (2011).

[29] McCright/Dunlap (2011); Munro/Ditto (1997); Lord/Ross/Lepper (1979).

[30] Cowen (2005).

[31] Bénabou/Tirole (2016).

[32] Korn et al. (2014).

[33] Winter (1993); Winter (2004); Winter (2007).

Maschinelles Lernen

Implizite Motive wurden bis jetzt im TAT oder vergleichbaren Testverfahren von Psychologen in zeitaufwendiger Handkodierung erfasst.[34] Die Fortschritte computergestützter Analyseverfahren ermöglichen es heute, geschriebenen Text automatisiert auszuwerten. Schultheiss hat auf der Basis des Linguistic Inquiry and Word Count (LIWC)[35] einen ersten Ansatz dazu vorgelegt.[36] Die Hypothese, dass es Marker-Worte für die Motivmessung gibt, konnte bestätigt werden.

Die Studie basiert allerdings auf einer nur kleinen Stichprobe für Deutschland und die USA, zudem war die interkulturelle Übereinstimmung nicht überzeugend. Die Daten, welche wir in unserer Studie verwenden, wurden von einem Maschinenlernmodell ausgewertet, das aus einer sehr viel größeren Stichprobe entwickelt wurde. Die von Johannßen et al. vorgestellte Methode nutzt ein überwachtes Lernmodell. Anders gesagt, lernt der Algorithmus, den Motivkategorien selbstständig Textmerkmale zuzuordnen.[37]

Zum Anlernen eines Algorithmus wurden zwei Trainigsdatensätze verwendet, die Expert:innen-Auswertungen nach dem Manual von Kuhl zum Operanten Motivtest (OMT)[38] umfassen; die Testpersonen schreiben jeweils zu 15 Bildern kurze Geschichten.[39]

Der auf den kodierten OMT-Antworten basierende Trainingsdatensatz wurde hinsichtlich der Interpunktion sowie von Funktionswörtern bereinigt. Die Vereinfachung der Sprache reduziert das „Rauschen" sowie die erforderliche Rechenleistung. Hierfür wurden die Texte in ihre logischen Einheiten (Tokens) zerlegt und Programmier- und Rechtschreibfehler in automatisierter Form eliminiert.[40]

Im folgenden Schritt wurden aus den Trainingsdaten Merkmale gewonnen, von denen ausgehend der Algorithmus die beste Vorhersagekraft entwickelt.[41] Die Auswertung erfolgt auf der Basis der Bag-of-Words-(BoW)Merkmals-

---

[34] Kuhl (2013); Winter (1991).

[35] Pennebaker/Francis/Booth (2001).

[36] Schultheiss (2013).

[37] Johannßen et al. (2020).

[38] Kuhl (2013).

[39] Johannßen/Biemann/Scheffer (2019); Johannßen et al. (2020).

[40] Zheng/Casari/Lotze (2019).

[41] Zheng/Casari/Lotze (2019).

gewinnung. Die BoW-Methode löst die Ordnung der Texttokens auf und gibt die Häufigkeit der Wörter innerhalb eines Textes an. Es werden Merkmalsvektoren generiert, die die Häufigkeit eines Wortes über alle Dokumente enthalten.[42]

Aus dem durch die Bag-of-Words-Merkmalsgewinnung entstandenen Wörterbuch wurde eine Reihe von Merkmalslisten unter Berücksichtigung der Schwellenwerte erstellt. Eine Merkmalsliste mit der Frequenz 10 enthält somit alle Inhaltswörter, die mindestens zehn Mal in den Trainingsdaten verwendet wurden. Um die Klassifikationsgüte einer jeden Merkmalsliste zu bestimmen, wurde als Basis ein Naive-Bayes-Klassifikator verwendet. Ziel war es, die Merkmalsliste zu ermitteln, die voraussichtlich die höchste Vorhersagekraft besitzt.[43]

Die vier folgenden Algorithmen wurden getestet: Multinomial Naive Bayes (MNB), Discriminative Multinomial Naive Bayes (DMNB), Support Vector Machine (SVM) und Random Forest. Die Zuordnung für den Beitrag erfolgt auf der Grundlage des DMNB. In dem überwachten Lernprozess können die Klassifizierung jeder Instanz und die korrekte Zuordnung überprüft werden. Anhand der beobachteten Alpha- und Beta-Fehler werden Maßzahlen zu Präzision, Akkuranz und Rücklauf sowie das F1-Maß (Validität) errechnet.[44] Um Fehler durch zufälliges Rauschen zu vermeiden, wurden mehrere statistisch unabhängige Datensätze simuliert, indem eine k-fache Kreuzvalidierung durchgeführt wurde.[45]

Für die Daten konnten sehr gute Ergebnisse bezüglich des F-Scores (0,72), der Akkuranz (73,41), der Präzision (0,73) und des Rücklaufs (0,73) erzielt werden.[46] Nach Winter werden die Postings in einzelne Fenstergrößen von 50 Wörtern zerlegt.[47] Jedes Fenster wird anschließend anhand des entwickelten Maschinenlernmodells nach den Motiven kodiert. Durch ein Aggregieren der Ergebnisse für jedes Fenster wird das dominante Motiv für jeden Post kodiert. Falls ein Fenster weniger als 50 Wörter enthielt, wurde dieses dem vorherigen Fenster hinzugefügt.

Algorithmen kodieren bei Tests wie dem OMT beinahe genauso zuverlässig, wie menschliche Experten. Bei den vier getesteten Algorithmen konnten immer F-Werte von 0,71 bis 0,73 erreicht werden. Das ist insbesondere deshalb beachtlich, weil der Algorithmus in seiner ursprünglichen Fassung nicht nur

---

[42] Ebenda.

[43] Johannßen et al. (2020).

[44] Davis/Goadrich (2006).

[45] Zheng/Casari/Lotze (2019).

[46] Johannßen et al. (2020).

[47] Winter (1991).

wie beispielsweise bei Winter vier Kategorien (drei Motive und Null)[48], sondern 16 Kategorien erfasst. Für den vorliegenden Beitrag werden nur die besagten vier Kategorien verwendet.

Hypothesenbildung

Im vorherigen Teil dieses Beitrags wurde dargestellt, dass die lange Zeit in der Public Choice vorherrschende Annahme, dass Wähler:innen ihre Stimme nach rationalen Kriterien vergeben und diejenigen Politiker:innen oder Parteien wählen, die ihren individuellen Nutzen maximieren, schon seit Längerem als veraltet gilt. In Studien konnte nachgewiesen werden, dass bestehende Meinungen auf der Grundlage von Ideologie, Werten und Erfahrungen die Verarbeitung neuer Informationen beeinflussen.[49] Zusätzlich unterstützt die Kommunikation zwischen Subjekten bestehende Meinungen und Gruppenidentifikationen.[50] Es ist daher anzunehmen, dass die Art und Weise der politischen Kommunikation einen Einfluss auf die Verarbeitung durch Wähler:innen hat. Motivmuster können als unbewusste Elemente der Kommunikation einen Beitrag zum Verständnis der Wirkung politischer Kommunikation auf Wähler:innen liefern sowie den potenziellen Wählerkreis zu ermitteln.

Hypothese 1

Motive richten, wählen und aktivieren das Verhalten.[51] Sie bewirken, dass wir uns auf bestimmte Anreize in unserer Umwelt besonders stark reagieren, auf eine zeit- und situationsübergreifende Art und Weise. Eine Person mit einem besonders ausgeprägten Motiv richtet ihr Verhalten danach aus, bestimmte Reize direkt zu suchen, um durch ihre Reaktion auf sie Befriedigung zu erreichen.

Im Lauf der letzten 50 Jahre wurden zahlreiche Hypothesen aufgestellt und Studien durchgeführt, die den Einfluss der drei Motive auf die Politik oder politiknahe Bereiche zum Gegenstand hatten. So haben Studien ergeben, dass Politiker:innen mit einem starken Bindungsmotiv in historischen Untersuchun-

---

[48] Winter (1991).

[49] Le Yaouanq (2018); Rabin/Schrag (1999); Sunstein (2001).

[50] Sausgruber/Tyran (2011).

[51] Schultheiss/Brunstein (2010).

gen häufiger in politische Skandale verwickelt waren, aber auch die Bereitschaft zeigten, in politischen Krisen Konzessionen zu machen, die kriegerische Auseinandersetzungen verhinderten.[52] Ein starkes Leistungsmotiv konnte hingegen häufig mit einer weniger effizienten Administration infolge von mangelnder Delegation und Frustration in Verbindung gebracht werden.[53] Charismatische Politiker:innen konnten in einer Studie der US-Präsidenten, von Washington bis Lyndon B. Johnson, mit einem starken Machtmotiv in Verbindung gebracht werden.[54] Existierende Untersuchungen deuten demnach auf ein eindeutiges Profilmuster politischer Akteur:innen hin.

Hypothese 1: Das aggregierte Machtmotiv im Facebook-Datensatz sollte signifikant stärker sein als in dem repräsentativ erhobenen Testdatensatz. Das Bindungsmotiv sollte signifikant schwächer sein.

Hypothese 2

Ein sehr starkes Machtmotiv wird häufig mit dem Bedürfnis zu absoluter Willensdurchsetzung und mit einem starken persönlichen Geltungsdrang in Zusammenhang gebracht. Im politischen Kontext kann der Zusammenhang zwischen einem Hang zum Autoritären und einem übersteigerten Machtmotiv in Kombination mit einem schwach ausgeprägten Bindungsmotiv beobachtet werden.[55] Im demokratischen Kontext taucht die Motivkombination deutlich seltener auf.[56]

Die AfD wird auf der Grundlage von ausgewerteten Reden, Medienbeiträgen sowie weiterer Kommunikation mit einer rechtsautoritären Ausrichtung assoziiert.[57] Beispiele finden sich unter anderem in Buchveröffentlichungen[58]

---

[52] Weinberger/Cotler/Fishman (2010); Winter (2010).

[53] Winter (2010).

[54] Winter (2010).

[55] McClelland (1961).

[56] McClelland (1999).

[57] Thieme (2019).

[58] Höcke/Hennig (2018).

und Reden der Parteivorsitzenden[59]. Häusler verortet die AfD in einem völkisch-autoritären Populismus.[60]

Die Verortung der AfD als autoritäre Partei sollte, der Theorie folgend, ein signifikant stärkeres Machtmotiv und ein schwächeres Bindungsmotiv in der Kommunikation sichtbar machen. Sowohl interne Machtkämpfe als auch die Auseinandersetzung mit politischen Gegnern unterstützen die Hypothese.

Hypothese 2: Die AfD sollte ein signifikant stärkeres Machtmotiv aufweisen als die weiteren Parteien im Datensatz. Das Bindungsmotiv sollte signifikant schwächer sein.

Hypothese 3

Die Kommunikation der Parteien unterscheidet sich nach politischer und parteiinterner Kommunikation. Die parteiinterne Kommunikation enthält vor allem positive und gemeinschaftliche Aspekte, aber auch Konflikte innerhalb der Partei. Die politische Kommunikation hingegen erfordert die Auseinandersetzung mit externen Themen sowie mit politischen Gegnern. Die Kommunikation unterscheidet sich grundsätzlich je nach Zielpublikum und Zielsetzung. Außer in Fällen starker innerparteilicher Konflikte und Machtkämpfe ist die parteiinterne Kommunikation eher durch Einigkeit und Fakten (Termine, Wahlerfolge) geprägt.

Die Auseinandersetzung mit dem politischen Gegner über die Positionierung erfordert eine kritischere Kommunikation. Im Gesamtspektrum der politischen Akteur:innen herrscht bei der politischen Kommunikation weniger Einheit.

Hypothese 3a: Die parteiinterne Kommunikation sollte signifikant schwächer mit dem Machtmotiv kodiert sein als die parteiexterne Kommunikation.

Für das Erstarken der AfD gibt es verschiedene Erklärungsansätze. Zwei dieser Theorien sind für die strategische Wähleransprache relevant. Eine prominente Theorie findet sich in der „Modernisierungsverliererhypothese".[61] Sie besagt, dass durch wirtschaftliche Ängste im Zuge der fortschreitenden Globalisierung rechtspopulistische Tendenzen gestärkt werden. Die Auswirkungen

---

[59] Die Welt (11.9.2016); Frankfurter Allgemeine Zeitung (17.5.2018).

[60] Häusler (2018b); Häusler (2018a).

[61] Lengfeld (2017); Rippl/Seipel (2018).

sind in Bevölkerungsgruppen mit geringerem Humankapital und in Peripheriegegenden stärker; aber auch ein Teil der Mittelschicht, welche sich Sorgen um ihren wirtschaftlichen Abstieg macht, sei betroffen. Eine weitere These ist die „Cultural Backlash"-Theorie.[62] In dieser These werden weniger wirtschaftliche denn sozialpsychologische Gründe als ursächlich angesehen. Zusammengefasst wird angenommen, dass sich ein Teil der Bevölkerung von dem liberalen Wandlungsprozess in seiner Identität und seinen Wertesystemen bedroht fühlt.[63]

Für beide Theorien finden sich Belege. Das durchschnittliche Einkommen der Wähler:innen der AfD und der Linken ist am niedrigsten, zudem ist die Mehrheit der AfD-Wähler:innen der Auffassung, ökonomisch benachteiligt zu sein. Hinzu kommen der geringer Anteil an Hochschulabsolvent:innen und ein überproportionaler Anteil an Wähler:innen aus den Neuen Bundesländern und ländlichen Gebieten. Diese Aspekte sind in ähnlicher Ausprägung bei den Wähler:innen der Linken zu finden.[64]

Für die Cultural-Backlash-These spricht spezifisch, dass sich AfD-Wähler:innen besonders um Zuwanderung, Kriminalität, das Ende des Friedens und ein Auseinanderbrechen der Gesellschaft sorgen.

Die Darstellung der Wähler:innen stimmt mit der Erklärung der Motivforschung eines übersteigerten Machtmotivs überein. Studien haben gezeigt, dass typische Inhalte des Machtmotivs, wie die Angst vor Einflussverlust, Statusverlustr oder dem Verlust gesellschaftlichen Ansehens, bei Personen mit einem niedrigen oder bedrohten sozioökonomischen Status besonders ausgeprägt sind.[65]

Während die AfD mit einem völkisch-autoritären Populismus und zuwanderungskritischen Äußerungen auftritt und die Wählerschaft in prekären oder bedrohten sozioökonomischen Verhältnissen anspricht, setzt die Linke einen Akzent auf sozialpolitische Strategien und Arbeitnehmerperspektiven. Gemeinsam haben die Parteien einen Wähler:innenkreis mit niedrigem oder (gefühlt) bedrohtem sozioökonomischen Status.

Hypothese 3 ist in Teilen bereits in Hypothese 2 enthalten; nachdem von einer eindeutigen Differenzierung der parteiinternen und politischen Kommunikation ausgegangen werden kann, ist davon auszugehen, dass im Gesamtdatensatz nicht der reine Effekt der politischen Kommunikation messbar ist. Die

---

[62] Norris/Inglehart (2016); Rippl/Seipel (2018).

[63] Norris/Inglehart (2016).

[64] Brenke/Kritikos (2017).

[65] McClelland (1999).

politische Kommunikation sollte für die Wähleransprache und die Wählerentscheidung eine größere Rolle spielen als die parteiinterne Kommunikation.

Hypothese 3b: Die Parteien AfD und Die Linke an den politischen Rändern sollten ein signifikant stärkeres Machtmotiv bei politischen Themen aufweisen als die übrigen Parteien. Das Bindungsmotiv sollte signifikant schwächer ausgeprägt sein.

Daten und Modell

Der Datensatz umfasst die Kommunikation der Parteien CDU, CSU, SPD, FDP, Bündnis90 / DIE GRÜNEN, DIE LINKE, AfD auf dem sozialen Medium Facebook über den Zeitraum 2013 bis 2017.

Tab. 1: Deskriptive Statistik des Facebook-Datensatzes und des Trainingsdatensatzes über die einzelnen Motive

TOPIC DUMMY	NULL	BINDUNG	LEISTUNG	MACHT	TOTAL
Intern	422	1736	3330	7540	13028
	(3,24)	(13,33)	(25,56)	(57,88)	(100,00)
Politisch	385	1978	3782	10748	16893
	(2,28)	(11,71)	(22,39)	(63,62)	(100,00)
Total	807	3714	7112	18288	29921
	(2,70)	(12,41)	(23,77)	(61,12)	(100,00)
Testdaten	7988	18974	38294	75275	140531
	(5,16)	(13,31)	(26,63)	(54,90)	(100,00)

Der Facebook-Datensatz ist nach parteiinterner und politischer Kommunikation unterteilt.

Tabelle 1 zeigt die Kodierung der Textmeldungen nach den drei Motiven sowie der Null (Leer-)Kodierung. Untergliedert ist die Darstellung nach dem Facebook-Datensatz und dem Trainingsdatensatz. Die Tabelle zeigt eine durchschnittlich höhere Machtkodierung im Facebook-Datensatz als im Trainingsdatensatz und eine niedrigere Bindungs- und Leistungskodierung. Innerhalb des Facebook-Datensatzes zeigt sich, dass die politische Kommunikation ein stärkeres Machtmotiv als die parteiinterne Kommunikation aufweist. Für die Interpretation ist zu beachten, dass die politische Dummy-Variable eine hohe Varianz an Themen aufweist.

Neben der thematischen Kodierung und der Motivkodierung werden das Datum sowie die geteilten Inhalte (Status, Video, Foto, Veranstaltungen, Musik, Notiz, Link) als Kontrollvariablen hinzugezogen. Wir nutzen ein logarithmisches

Modell, um den Effekt der Parteizugehörigkeit in verschiedenen Spezifikationen der inkludierten Textmeldungen auf die Macht- beziehungsweise die Bindungskodierung zu untersuchen.

$$P(y_i = 1) = F(\beta_0 + \sum_{i=1}^{k} \beta_i X_i + \sum_{j=1}^{n} \alpha_j Z_j)$$

X_i repräsentiert die Textnachrichten der Parteien *i*, Z_i steht für die Kontrollvariablen des geteilten Contents *j*. Wir erstellen drei verschiedene Modelle. Zunächst wird nur die Parteizugehörigkeit geprüft, im nächsten Schritt werden die Kontrollvariablen mit einbezogen und im dritten Modell wird der Datensatz von Ausreißern bereinigt. Ausreißer sind in diesem Zusammenhang wie folgt zu verstehen: Es werden einzelne einflussreiche Kovariaten-Muster entfernt.[66] Auf diese Weise kann kontrolliert werden, ob Unterschiede durch einzelne Ereignisse bedingt sind oder über den Datensatz hinweg auftreten. Die Ergebnisse werden als durchschnittliche marginale Effekte (AME) angegeben.

Für die Darstellung des zeitlichen Verlaufs wird der „locally weighted scatterplot smoother" (LOWESS) genutzt.[67] Aufgrund der hohen Fallzahl wird ein niedriger Glättungsparameter von 0,3 angewandt.

Auswertung

Die Annahme, dass politische Akteur:innen das Motivmuster „starke Macht- und schwache Bindungskodierung" aufweisen, bestätigt sich (Tabelle 2). Das Modell bestätigt, dass die Kommunikation über Soziale Medien den zu erwartenden Mustern aus herkömmlicher Kommunikation ähnelt, also Politiker:innen ein tendenziell höheres Machtmotiv aufweisen.

Ebenfalls bestätigt sich die Annahme, dass die Parteien an den Rändern eine signifikant stärkere Machtkodierung aufweisen (Tabelle 3). Dabei bleibt sowohl der Unterschied zwischen der Linkspartei und den Parteien CDU, CSU, FDP, SPD sowie BÜNDNIS 90 / DIE GRÜNEN Grünen als auch der Unterschied zwischen der Linkspartei und der AfD in die entgegengesetzte Richtung hochsignifikant. Die Ergebnisse sind gegenüber der Einführung der Kontrollvariablen ebenso wie der Eliminierung der Ausreißer robust. Die signifikant erhöhten

[66] Kohler/Kreuter (2017).

[67] Cleveland (1979).

Machtkodierungen und geringere Bindungskodierungen der AfD bestätigen, dass sich die rechtspopulistische Orientierung nicht nur in der thematischen Schwerpunktsetzung, sondern auch in der Motivkodierung zeigt.

Tab. 2: Logistisches Modell des Macht- und Bindungsmotivs des Testdatensatzes und Facebook-Datensatzes

	MODELL 1 MACHT-MOTIV	MODELL 2 BINDUNGS-MOTIV
Testdatensatz	0	0
Facebook-Daten	0,0756***	-0,0109***
Beobachtungen	170452	170452

$^{*}p < 0{,}05$, $^{**}p < 0{,}01$, $^{***}p < 0{,}001$

Tab. 3: Logistisches Modell der Macht- und Bindungsmotive nach Parteien

	MODELL 3 MACHT-MOTIV	MODELL 4 MACHT-MOTIV	MODELL 5 MACHT-MOTIV	MODELL 6 BILDUNGS-MOTIV	MODELL 7 BILDUNGS-MOTIV	MODELL 8 BILDUNGS-MOTIV
AfD	0,0265**	0,0437***	0,0562***	-0,0157**	-0,0274***	-0,0224***
CDU	-0,0611***	-0,0651***	-0,0659***	0,0289***	0,0302***	0,0262***
CSU	-0,0874***	-0,0766***	-0,0818***	0,0553***	0,0450***	0,0357***
FDP	-0,0358***	-0,0388***	-0,0225*	0	0,000405	0
BÜNDNIS 90 / DIE GRÜNEN	-0,0776***	-0,0765***	-0,0847***	0,0483***	0,0460***	0,0413***
DIE LINKE	0	0	0	0,00303	0	0,00289
SPD	-0,0417***	-0,0395***	-0,0355***	0,0321***	0,0284***	0,0313***
Datum		-0,0000537***	-0,0000565***		0,0000370***	0,0000384***
Kontrollvariablen		x	x		x	x
Kovariaten-Muster			x			x
Beobachtungen	29921	29921	28674	29921	29917	28004

$^{*}p < 0{,}05$, $^{**}p < 0{,}01$, $^{***}p < 0{,}001$

Im nächsten Schritt werden die Ergebnisse der Motivanalyse mit der inhaltsanalytischen Auswertung kombiniert. Es wird zwischen inhaltlich, politischen Fragestellungen und der parteiinternen Kommunikation unterschieden.

Tab. 4: Logistisches Modell der Machtkodierung nach Parteien – Unterschiede der innerparteilichen und politischen Kommunikation

	MODELL 9 MACHT-MOTIV AFD	MODELL 10 MACHT-MOTIV CDU	MODELL 11 MACHT-MOTIV CSU	MODELL 12 MACHT-MOTIV FDP	MODELL 13 MACHT-MOTIV GRÜNE	MODELL 14 MACHT-MOTIV LINKE	MODELL 15 MACHT-MOTIV SPD
TopicDummy	-0,0369*	0,00816	-0,121***	-0,0270*	-0,0704***	-0,0286*	-0,0461**
Datum	-0,00000815	-0,00000724	-0,000106***	-0,0000887***	-0,0000406*	-0,0000571***	-0,0000774***
Kontroll-variabelen	x	x	x	x	x	x	x
Beobachtungen	4908	3432	3423	5615	3246	5113	4174

** p < 0,05, ** p < 0,01, *** p < 0,001*
Anmerkung: Interne Kommunikation ist mit 1 kodiert.

Die Annahme, dass die Parteien intern eine weniger stark machtkodierte Kommunikation praktizieren, konnte für nahezu alle Parteien bestätigt werden. Lediglich die CDU weist weder eine signifikante Unterscheidung zwischen parteiinterner und politischer Kommunikation noch den antizipierten negativen Zusammenhang auf. Für alle weiteren Parteien ist der erwartete negative Zusammenhang erkennbar und signifikant.

Tab. 5: Logistisches Modell der Machtkodierung nach Parteien – innerparteiliche Kommunikation

	MODELL 16 MACHT-MOTIV INTERN	MODELL 17 MACHT-MOTIV INTERN	MODELL 18 MACHT-MOTIV INTERN
AfD	0,015	0,033	0,0427*
CDU	-0,0369*	-0,0412**	-0,0481**
CSU	-0,119***	-0,103***	-0,111***
FDP	-0,0310*	-0,0313*	-0,00768
BÜNDNIS 90 / DIE GRÜNEN	-0,0989***	-0,0944***	-0,0992***
DIE LINKE	0	0	0
SPD	-0,0552**	-0,0459**	-0,0512**
Datum		-0,0000513***	-0,0000495***
Kontrollvariablen		x	x
Kovariaten-Muster			x
Beobachtungen	13335	13335	12879

** p < 0,05, ** p < 0,01, *** p < 0,001*

Tab. 6: Logistisches Modell der Machtkodierung nach Parteien – politische Kommunikation

	MODELL 19 MACHT-MOTIV POLITISCH	MODELL 20 MACHT-MOTIV POLITISCH	MODELL 21 MACHT-MOTIV POLITISCH
AfD	0,0245*	0,0388**	0,0513***
CDU	-0,0785***	-0,0834***	-0,0665***
CSU	-0,0291	-0,0225	-0,0173
FDP	-0,023	-0,0265	-0,009
BÜNDNIS 90 / DIE GRÜNEN	-0,0555***	-0,0588***	-0,0455**
DIE LINKE	0	0	0
SPD	-0,0366**	-0,0401**	-0,0269*
Datum		-0,0000596***	-0,0000605***
Kontrollvariablen		x	x
Kovariaten-Muster			x
Beobachtungen	16586	16584	16095

** p < 0,05, ** p < 0,01, *** p < 0,001*

Basierend auf den Theorien des Cultural Backlash und der Modernisierungsverlierer wurde die Hypothese formuliert, dass die Parteien an den politischen Rändern mit einem erhöhten Wählerpotenzial aus prekären Verhältnissen oder peripheren Gebieten ebendieses Wählerpotenzial mit einer erhöhten Machtkodierung und geringeren Bindungskodierung ansprechen.

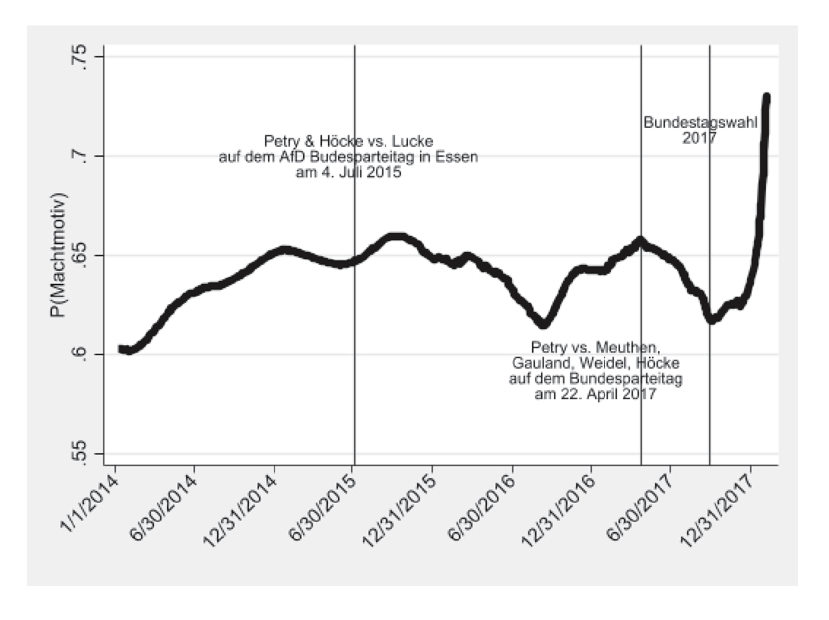
Die Hypothese, dass Parteien an den politischen Rändern ein stärkeres Motivmuster in der politischen Kommunikation aufweisen, kann für die AfD eindeutig bestätigt werden (Tabelle 6). Wir sehen eine signifikant stärkere Machtkodierung im Vergleich mit der nächstliegenden Partei DIE LINKE. Letztere weist den zweithöchsten Wert auf, unterscheidet sich dabei jedoch nicht signifikant von der FDP und der CSU. Die Bindungskodierung ist weniger robust gegenüber den Modellspezifikationen. Zwar ist die Bindungskodierung der AfD am schwächsten, sie unterscheidet sich aber nicht signifikant von dem Basiswert der FDP. Signifikant ist hingegen der Unterschied im Vergleich mit den Grünen und der SPD sowie der Linken in der dritten Spezifikation. Insgesamt ist die Bindungskodierung des linken Parteienspektrums höher als des rechten Parteienspektrums.

Auf die Hypothese bezogen kann es als bestätigt angesehen werden, dass die AfD mit einer ultrakonservativen bis völkisch-populistischen Ausrichtung eine erhöhte Machtkodierung aufweist (Tabelle 6). Dazu passend ist, dass die

Wählerschaft, der psychologischen Literatur folgend, für diese Motivkodierung anfällig ist.

In der parteiinternen Kommunikation (Tabelle 5) ist die Kodierung weniger eindeutig. Es zeigt sich allerdings, dass die Regierungsparteien und die Grünen mit einer stabilen Führungsspitze über den Zeitraum ein signifikant geringer kodiertes Machtmotiv aufweisen. Die AfD mit dem im Beobachtungszeitraum liegenden Flügelkampf weist die höchste Machtkodierung auf. Die Bindungskodierung ist den Modellspezifikationen gegenüber nicht robust.

Über den betrachteten Untersuchungszeitraum konnte für die AfD nachgewiesen werden, dass die Machtkodierung in Zeiten innerparteilicher Konflikte ansteigt. Die Daten ermöglichen es, einzelne Events im Zeitverlauf nicht nur auf der Basis inhaltlicher Aspekte, sondern auch durch die Kodierung impliziter Motive nachzuweisen. In Abbildung 1 sind der Anstieg der Machtkodierung bis zum Parteitag am 4. Juli 2015 und der Anstieg hin zum Bundesparteitag am 22. April 2017 sowie ein deutlicher Rückgang der Machtkodierung hin zur Bundestagswahl 2017 zu erkennen.



Anmerkung: Glättungsfaktor 0,3.

Abb. 1: LOWESS-Plot – Verlauf der Machtkodierung der innerparteilichen Kommunikation für Beobachtungen der AfD

Tab. 7: Logistisches Modell Machtkodierung nach Parteien – Migration im Vergleich mit der übrigen politischen Kommunikation

	MODELL 22 MACHT-MOTIV AFD	MODELL 23 MACHT-MOTIV CDU	MODELL 24 MACHT-MOTIV CSU	MODELL 25 MACHT-MOTIV FDP	MODELL 26 MACHT-MOTIV GRÜNE	MODELL 27 MACHT-MOTIV LINKE	MODELL 28 MACHT-MOTIV SPD
Migr_ Dummy	-0,0467**	0,179**	0,0590*	-0,0979**	-0,0241	0,0325	-0,0262
Datum	0,0000012	-0,00000395	-0,000103**	-0,000105***	-0,0000721**	-0,0000493*	-0,0000951***
Kontroll-variablen	x	x	x	x	x	x	x
Beobach-tungen	3747	1466	1444	2141	1786	3221	2775

** p < 0,05, ** p < 0,01, *** p < 0,001*
Anmerkung: Datenbasis politische Kommunikation, Migration ist mit 1 kodiert.

Neben der groben Unterscheidung zwischen parteiinterner und politischer Kommunikation wurde das relevanteste Thema über die Legislaturperiode für eine genauere Betrachtung herausgefiltert. Das Migrationsthema ist zudem als Hauptthema der AfD ab Mitte 2015 anzusehen. Im Vergleich mit den weiteren politischen Themen ist bei der CDU sowie der CSU eine signifikant stärkere Machtkodierung festzustellen. Die AfD hingegen weist eine geringere Machtkodierung als in den weiteren politischen Themen auf.

Die AfD ist jedoch die einzige Partei mit einem ansteigenden Machtmotiv im Bereich Migration im Zeitverlauf, wenngleich das Ergebnis nicht signifikant ist. Die weiteren Parteien weisen bis auf die CDU einen signifikant negativen Zusammenhang mit der Zeitvariablen auf. Dies deutet darauf hin, dass bei allen Parteien außer der AfD das Thema nach einem Hoch im Jahr 2015 im Rahmen der Bundestagswahl weniger Bedeutung hatte.

Schlussfolgerungen

Es kann zunächst festgestellt werden, dass die Machtkodierung politischer Kommunikation auf Sozialen Medien stärker ausgeprägt ist als im Vergleichsdatensatz. Unterstützt wird dieses Ergebnis von der Annahme, dass Politiker:innen grundsätzlich ein verstärktes Machtmotiv besitzen.

Ein Motiv alleine wird den Wahlerfolg einer Partei nicht determinieren. Existierende Studien legen nahe, dass in den USA jene Präsidentschaftskandidaten gewählt wurden, die dem Motivmuster der Bevölkerung am nächsten

kamen.[68] Erste Ergebnisse auf der Basis diesen Datensatzes liefern ähnliche Zusammenhänge wie in dieser Theorie. Aus dem Ergebnis ließe sich ableiten, dass implizite Inhalte bestimmte Wählergruppen besonders ansprechen. In Zukunft könnte die Untersuchung impliziter Motive daher einen Anhaltspunkt für die Prognose von Wahlergebnissen bieten.

In den Untersuchungen zu Motivmustern der Sozialen Medien konnten die signifikanten Unterschiede zwischen den Parteien herausgearbeitet werden. Insbesondere weist die Polarisierung der Parteienlandschaft und die damit einhergehende signifikant stärkere Machtkodierung / geringere Bindungskodierung der AfD in Kombination mit einem wachsenden Wahlerfolg im Beobachtungszeitraum darauf hin, dass sich die Wählerschaft dieser Motivkodierung zugehörig fühlt.

Zudem kann ein signifikanter Unterschied zwischen der politischen und der parteiinternen Kommunikation festgestellt werden, der robust negativ signifikant für alle Parteien außer der CDU ist. Insbesondere bei politischen Beiträgen der CDU fällt ein hohes Maß an Darstellung politischer Erfolge auf. Dies würde erklären, weshalb ein umgekehrter Zusammenhang beobachtbar ist. Dieser wiederum könnte in der Position als „großer" Partner in der großen Koalition und Partei der Bundeskanzlerin begründet liegen.

Die tiefergehende Analyse einzelner politischer Themenfelder konnte am Beispiel der Migrationspolitik interessante Ergebnisse liefern. Es ist ein zunächst paradoxes Muster zu beobachten. Entgegen der intuitiven Vermutung, dass die AfD das stärkere Machtmotiv in ihrem Hauptthema besitzen sollte, ist das Machtmotiv signifikant geringer kodiert als im Fall der übrigen politischen Themen. Umgekehrtes ist bei der CDU und der CSU zu beobachten: Die Machtkodierung ist robust über die Modellspezifikationen 17 bis 19 % höher als bei den weiteren politischen Themen. Eine mögliche Erklärung ist in der Konzeption der Motivmuster zu finden. Diese sind als Bedürfnisse zu verstehen. Wenn in einem Soll-Ist-Vergleich der Ist-Wert den Soll-Wert übersteigt, gibt es keine Diskrepanz mehr.[69] Im Fall der AfD gibt es durch die Erfolge in der Bestimmung des öffentlichen Diskurses (von „Willkommenskultur" hin zu Abschottung) und eine reale Reduzierung der Zuwanderung kaum nennenswerte Soll-Ist-Diskrepanzen. CDU und CSU hingegen waren in der Legislaturperiode (2013 bis 2017) getrieben von der AfD und der eigenen Basis. Das kann als eine Bedrohung der Stellung im

---

[68] Winter (2010).

[69] McClelland (1999).

konservativen Spektrum und damit ein Auseinanderfallen von Anspruch und Wirklichkeit verstanden werden. Dies sollte zu einer heftigen emotionalen, auf stammesgeschichtlich alten Motiven beruhenden Reaktion führen, die einer rationalen Kosten-Nutzen-Erwägung vorgelagert ist.

Der extreme Anstieg des Machtmotivs beim Thema Migration bei CDU und CSU weist also auf ein emotionales Defizit hin: Es wurde wenig über das Thema gesprochen, dann jedoch emotional umso heftiger, mit einer hohen Machtkodierung, um sich im politischen Wettbewerb zu behaupten. Diese Kommunikationsform kann unausgewogen wirken, das kann auf die emotionale und impulsive Natur von Motiven, insbesondere des Machtmotivs zurückgeführt werden.

Der Beitrag zeigt, dass die Auswertung hinsichtlich psychologischer Motive in Kombination mit inhaltsanalytischen Methoden zu interpretierbaren Ergebnissen führt. Da implizite Motive unwillkürlich sind, stellen die Ergebnisse eine vielversprechende Alternative zu Sentiment Scores in der politischen Analyse dar. Die Kombination der Techniken erlaubt eine tiefergehende Analyse informeller Kommunikation mit einem wenig gesteigerten Zeitaufwand. Gleichwohl bleiben die Ergebnisse immer sensibel gegenüber den Annahmen der zugrunde liegenden Algorithmen.

Literatur

Bénabou, Roland / Tirole, Jean (2016): Mindful economics: The production, consumption, and value of beliefs. In: Journal of Economic Perspectives, Vol. 30, Nr. 3, S. 141–164.

Brenke, Karl / Kritikos, Alexander S. (2017): Wählerstruktur im Wandel, in: DIW Wochenbericht, Nr. 29, S. 595–606.

Caplan, Bryan (2008): The Myth of the Rational Voter. Why Democracies Choose Bad Policies, Neuauflage, Princeton University Press, Princeton und Oxford.

Cleveland, William S. (1979): Robust locally weighted regression and smoothing scatterplots. In: Journal of the American Statistical Association, Vol. 74, Nr. 368, S. 829–836.

Cowen, Tyler (2005): Self-deception as the root of political failure. In: Public Choice,-Vol. 124, Nr. 3–4, S. 437–451.

Davis, Jesse / Goadrich, Mark (2006): The relationship between Precision-Recall and ROC curves. In: Cohen, William und Andrew Moore (Hrsg.): Proceedings of the 23rd international conference on Machine learning, S. 233–240.

Die Welt (2016): AfD: Frauke Petry will Begriff „völkisch“ positiv besetzen, 11.9.2016.

Downs, Anthony (1997): An Economic Theory of Democracy, Paperback, 1997; Originalausgabe erschienen 1957 bei Harper, New York.

Fodor, Eugene M. (2010): Power Motivation. In: Implicit Motives. Oxford University Press, Oxford, S. 3–29.

Frankfurter Allgemeine Zeitung (2018): „In Anatolien entsorgen“: Ermittlungen gegen Gauland eingestellt, 17.5.2018.

Häusler, Alexander (2018a): Die AfD: Partei des völkisch-autoritären Populismus. In: Häusler, Alexander (Hrsg.): Völkisch-autoritärer Populismus. Der Rechtsruck in Deutschland und die AfD. VSA: Verlag Hamburg. Hamburg, S. 9–19.

Häusler, Alexander (2018b): Einleitung. In: Häusler, Alexander (Hrsg.): Völkisch-autoritärer Populismus. Der Rechtsruck in Deutschland und die AfD. VSA: Verlag Hamburg. Hamburg, S. 7f.

Herz, Dietmar / Weinberger, Veronika (Hrsg; 2019): Die 100 wichtigsten Werke der Ökonomie. Von Adam Smith bis Paul Krugman. Schäffer-Poeschel Verlag, Stuttgart.

Höcke, Björn / Hennig, Sebastian (2018): Nie zweimal in denselben Fluß: Björn Höcke im Gespräch mit Sebastian Hennig.

Johannßen, Dirk / Biemann, Chris / Scheffer, David (2019): Reviving a psychometric measure: Classification and prediction of the Operant Motive Test. In: Niederhoffer, K. / Hollingshead, K. / Resnik, P. / Resnik, R. / Loveys, K. (Hrsg.): Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology, Stroudsburg, S. 121 -125.

Johannßen, Dirk / van Heteren-Fresen, Nicolas / Scheffer, David / Islam, Zahurul / Biemann, Chris (2020): Creation of an automatized and accurate, stable and valid AI model for the prediction of implicit motives - unveröffentlicht.

Kahneman, Daniel (2003): A perspective on judgment and choice: mapping bounded rationality. In: American Psychologist,Vol. 58, Nr. 9, S. 697.

Kohler, Ulrich / Kreuter, Frauke (2017): Datenanalyse mit Stata. Allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung. De Gruyter, Oldenbourg, Berlin.

Korn, Christoph W. / Sharot, Tali / Walter, Hendrik / Heekeren, Hauke R. / Dolan, Raymond J. (2014): Depression is related to an absence of optimistically biased belief updating about future life events. In: Psychological medicine,Vol. 44, Nr. 3, S. 579–592.

Kuhl, Julius (2013): Auswertungsmanual für den Operanten Multi-Motiv-Test OMT: basierend auf Julius Kuhl & David Scheffer (2009). Sonderpunkt Wissenschaftsverlag Münster, Greven.

Le Yaouanq, Yves (2018): A model of ideological thinking. Discussion Paper No. 85, Ludwig-Maximilians-Universität München.

Lengfeld, Holger (2017): Die „Alternative für Deutschland „: eine Partei für Modernisierungsverlierer? in: KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie,-Vol. 69, Nr. 2, S. 209–232.

Lord, Charles G. / Ross, Lee / Lepper, Mark R. (1979): Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. In: Journal of personality and social psychology,Vol. 37, Nr. 11, S. 2098.

McClelland, David C. (1961): The Achieving society. Van Nostrand, Princeton.

McClelland, David C. (1999): Human motivation. Cambridge University Press, 6. Aufl., Cambridge.

McClelland, David C. / Atkinson, John W. (1948): The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. In: The Journal of Psychology,Vol. 25, Nr. 2, S. 205–222.

McCright, Aaron M. / Dunlap, Riley E. (2011): The politicization of climate change and polarization in the American public's views of global warming, 2001–2010. In: The Sociological Quarterly,Vol. 52, Nr. 2, S. 155–194.

Munro, Geoffrey D. / Ditto, Peter H. (1997): Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. In: Personality and Social Psychology Bulletin,Vol. 23, Nr. 6, S. 636–653.

Murray, Henry Alexander (1943): Thematic apperception test. Harvard University Press, Cambridge, MA.

Norris, Pippa / Inglehart, Ronald (2016): Trump, Brexit, and the rise of populism: Economic have-nots and cultural backlash. In: Harvard JFK School of Government Faculty Working Papers Series, No. RWP16-0269, S. 1–52.

Pang, Joyce (2010): The Achievement Motive: A Review of Theory and Assessment of N Achievement, Hope of Success, and Fear of Failure. In: Implicit motives. Oxford University Press, Oxford, S. 30–70.

Pennebaker, James / Francis, Martha / Booth, Roger (2001): Linguistics inquiry and word count (LIWC). Online verfübar unter https://www.researchgate.net/publication/246699633_Linguistic_inquiry_and_word_count_LIWC.

Rabin, Matthew / Schrag, Joel L. (1999): First impressions matter: A model of confirmatory bias. In: The quarterly journal of economics,Vol. 114, Nr. 1, S. 37–82.

Rippl, Susanne / Seipel, Christian (2018): Modernisierungsverlierer, Cultural Backlash, Postdemokratie. In: KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie,Vol. 70, Nr. 2, S. 237–254.

Sausgruber, Rupert / Tyran, Jean-Robert (2011): Are we taxing ourselves?: How deliberation and experience shape voting on taxes. In: Journal of Public Economics,-Vol. 95, Nr.1–2, S. 164–176.

Schnellenbach, Jan / Schubert, Christian (2015): Behavioral political economy: A survey. In: European Journal of Political Economy,Vol. 40 (2015), S. 395–417.

Schultheiss, Oliver C. (2013): Are implicit motives revealed in mere words? Testing the marker-word hypothesis with computer-based text analysis. In: Frontiers in psychology,Vol. 4, S. 748.

Schultheiss, Oliver C. / Brunstein, Joachim C. (2010): Introduction. In: Brunstein, Joachim C. und Schultheiss, Oliver C. (Hrsg.): Implicit motives. Oxford University Press, Oxford, S. ix–xxvii.

Shipley Jr, Thomas E. / Veroff, Joseph (1952): A projective measure of need for affiliation. In: Journal of experimental psychology,Vol. 43, Nr. 5, S. 349–356.

Simon, Hermann / Fassnacht, Martin (1982): Preismanagementt. Springer Gabler, Wiesbaden.

Steele, Robert S. (1977): Power motivation, activation, and inspirational speeches. In: Journal of Personality, Nr. 45, S. 53–64.

Sunstein, Cass R. (2001): Republic.com. Princeton University Press, Princeton.

Thieme, Tom (2019): Dialog oder Ausgrenzung – Ist die AfD eine rechtsextreme Partei? Online abrufbar unter https://www.bpb.de/politik/extremismus/rechtspopulismus/284482/dialog-oder-ausgrenzung-ist-die-AfD-eine-rechtsextreme-partei, zuletzt aufgerufen am 4.1.2021.

Tversky, Amos / Kahneman, Daniel (1986): Rational Choice and the Framing of Decisions. In: The Journal of Business,Vol. 59, Nr. 4, S. 251–278.

Weinberger, Joeal / Cotler, Tanya, Fishman, Daniel (2010): The Duality of Affiliative Motivation. In: Implicit motives. hrsg. von Brunstein, J. C. / Schultheiss, O. C. 2010, S. 71–88.

Winter, David G. (1991): Measuring personality at a distance: Development of an integrated system for scoring motives in running text. In: Perspectives in personality. Vol. 3, S. 59–89.

Winter, David G. (1993): Power, affiliation, and war: Three tests of a motivational model. In: Journal of personality and social psychology,Vol. 65, Nr. 3, S. 532–545.

Winter, David G. (2004): Motivation and the escalation of conflict: Case studies of individual leaders. In: Peace and Conflict: Journal of Peace Psychology, Vol. 10, Nr. 4, S. 381–398.

Winter, David G. (2007): The role of motivation, responsibility, and integrative complexity in crisis escalation: comparative studies of war and peace crises. In: Journal of personality and social psychology, Vol. 92, Nr. 5 (2007), S. 920–937.

Winter, David G. (2010): Political and Historical Consequences of Implicit Motives. In: Brunstein, Joachim C. und Schultheiss, Oliver C. (Hrsg.): Implicit motives. Oxford University Press, Oxford, S. 407–432.

Zheng, Alice / Casari, Amanda (2019): Merkmalskonstruktion für machine learning. Prinzipien und Techniken der Datenaufbereitung, dpunkt.verlag, Heidelberg.

Verfasserinnen und Verfasser

Islam , Zahurul, Professor an der NORDAKADEMIE Hochschule der Wirtschaft

Karaman Örsal, Deniz Dilan, Dr. rer. pol. (Humboldt-Universität zu Berlin), Universität Hamburg und außerplanmäßige Professorin an der Leuphana Universität Lüneburg.

Maass, Christina Heike, M. Sc. in Economics, Universität Hamburg

Roth, Felix, Privatdozent für Volkswirtschaftslehre an der Universität Hamburg und Leiter des Projekts GLOBALINTO im Rahmen der Horizon-2020-Forschungsförderung der Europäischen Kommission

Scheffer, Niklas, cand. rer. pol. (Universität Potsdam), Universität Hamburg, Institut für Computer Aided Psychometric Text Analysis (CAPTA)

Schneider, Henrique, Professor für Volkswirtschaftslehre an der Nordakademie, Hochschule der Wirtschaft, in Elmshorn und stellvertretender Direktor des Schweizerischen Gewerbeverbands sgv in Bern, Schweiz

Straubhaar, Thomas, Professor für Volkswirtschaftslehre, insbesondere Internationale Wirtschaftsbeziehungen der Universität Hamburg

Sturm, Silke, M. Sc (Universität Bayreuth), Universität Hamburg

Vöpel, Henning, Hamburgisches WeltWirtschaftsinstitut (HWWI) und Professor der Hamburg School of Business Administration (HSBA)



Reihe Edition HWWI
herausgegeben von Thomas Straubhaar

In der Edition HWWI (ISSN 1865-7974) erscheinen abgeschlossene, umfangreiche Projektergebnisse sowie Dissertationen zu Forschungsthemen, die vom HWWI bearbeitet werden . Folgende Titel sind bisher erschienen:

Band 1: Thomas Straubhaar (Hg.): Bedingungsloses Grundeinkommen und Solidarisches Bürgergeld – mehr als sozialutopische Konzepte, 2008.
ISBN 978-3-937816-47-0, DOI https://doi.org/10.15460/HUP.HWWI.1.69.

Band 2: Martin-Peter Büch et al. (Hg.): Sportfinanzierung – Spannungen zwischen Markt und Staat, 2009.
ISBN 978-3-937816-53-1, DOI https://doi.org/10.15460/HUP.HWWI.2.70.

Band 3: Martin-Peter Büch et al. (Hg.): Zur Ökonomik von Spitzenleistungen im internationalen Sport, 2012.
ISBN 978-3-937816-87-6, DOI https://doi.org/10.15460/HUP.HWWI.3.122.

Band 4 : Martin-Peter Büch et al. (Hg.): Sport und Sportgroßveranstaltungen in Europa – zwischen Zentralstaat und Regionen, 2012.
ISBN 978-3-937816-88-3, DOI https://doi.org/10.15460/HUP.HWWI.4.123.

Band 5: Seçil Paçacı Elitok, Thomas Straubhaar (eds.): Turkey, Migration and the EU: Potentials, Challenges and Opportunities, 2012.
ISBN 978-3-937816-94-4, DOI https://doi.org/10.15460/HUP.HWWI.5.118.

Band 6: Thomas Straubhaar (Hg.): Neuvermessung der Datenökonomie, 2021.
ISBN (Print) 978-3-943423-91-4, (Epub) 978-3-943423-94-5,
DOI https://doi.org/10.15460/HUP.HWWI.6.212.

Die Online-Ausgaben der Reihe sind frei zugänglich als Open-Access-Publikation erschienen. Die Printversion kann über den Buchhandel oder direkt beim Verlag (https://hup.sub.uni-hamburg.de) bezogen werden.